



Vers un système d'évaluation de la qualité d'image multi-critères

Aladine Chetouani

► To cite this version:

Aladine Chetouani. Vers un système d'évaluation de la qualité d'image multi-critères. Traitement du signal et de l'image [eess.SP]. paris; Université Paris-Nord - Paris XIII, 2010. Français. NNT : . tel-00560808

HAL Id: tel-00560808

<https://theses.hal.science/tel-00560808>

Submitted on 2 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT GALILEE - UNIVERSITÉ PARIS 13

THESE

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ PARIS 13

présentée et soutenue publiquement :

Aladine Chetouani

Titre :

**Vers un système d'évaluation de la
qualité d'image multi-critères**

Soutenue le 29 Novembre 2010 devant le jury composé de :

P	Patrick Viaris De Lesegno	Président	L2TI, Université Paris 13
P.	Patrick Le Callet	Rapporteur	IVC, Polytech Nantes
P.	Alain Trémeau	Rapporteur	LIGIV, Université Jean-Monnet
P.	Mohamed Deriche	Examineur	KFUPM (KSA)
P.	Jenny Benois-Pineau	Examineur	LaBRI, Université Bordeaux 1
P.	Azeddine Beghdadi	Directeur de thèse	L2TI, Université Paris 13
Mcf.	Ghilès Mostafaoui	Co-encadrant	ETIS, Université de Cergy

Remerciements

Je tiens avant tout à remercier mon directeur de thèse Monsieur Azeddine Beghdadi, directeur du laboratoire L2TI (Laboratoire de Traitement et de Transport de l'Information) de m'avoir permis d'effectuer une thèse, ainsi que mon co-encadrant de thèse Monsieur Ghilès Mostafaoui. Je les remercie pour leur encadrement, leurs conseils, et leur disponibilité.

Je souhaite également à remercier l'ensemble des membres du jury qui m'ont fait l'honneur de lire ce manuscrit et de commenter mes travaux de recherche : Messieurs Patrick Le Callet et Alain Trémeau pour avoir rapporté ce manuscrit ainsi que pour la qualité et la pertinence de leurs remarques, Monsieur Patrick Viaris De Lesegno pour avoir accepté de présider le jury et enfin Madame Jenny Benois-Pineau et Monsieur Mohamed Deriche, pour leur expertise et leurs remarques pertinentes. Je tiens également à remercier l'ensemble des membres du laboratoire, permanents, thésards et personnels pour m'avoir accepté parmi eux pendant ces trois années.

Un grand merci à ma famille qui m'a apporté le soutien nécessaire durant les périodes difficiles. Merci pour leur patience et leurs encouragements. Merci à tous et particulièrement mes parents, grâce à qui j'ai pu atteindre ce stade.

Ce travail a pu être réalisé en partie grâce au projet HD3D.iio, il convient donc de remercier également l'ensemble des personnes s'inscrivant dans ce projet.

Table des matières

Table des figures	iv
Liste des tableaux	x
Introduction générale	3
0.1 Contexte	5
0.2 Organisation de la thèse	6
0.3 Liste des publications et communications	8
1 La qualité d'image	11
1.1 Introduction	13
1.2 Notions et définitions	17
1.3 Evaluation subjective	18
1.3.1 Les facteurs influant	18
1.3.2 Les protocoles	20
1.3.3 Bases d'images dédiées à l'évaluation de performance	24
1.4 Evaluation objective	28
1.4.1 Approches basées EQM	30
1.4.2 Approches structurelles	33
1.4.3 Approches inspirées du Système Visuel Humain	35
1.4.4 Approches neuronales	58
1.5 Conclusion	59
I Etudes des dégradations	63
2 Etude des Effets de bloc	65
2.1 Introduction	66

2.2	Etat de l'art	67
2.2.1	Approches spatiales	67
2.2.2	Approches fréquentielles	69
2.3	Prédiction des effets de blocs	70
2.3.1	Description de la méthode	71
2.3.2	Résultats expérimentaux	76
2.3.3	Application : Tatouage	78
2.4	Estimation de la visibilité des effets de bloc	83
2.4.1	Carte de visibilité	83
2.4.2	Réduction des effets de bloc	88
2.5	Conclusions	93
3	Estimation du flou	95
3.1	Introduction	96
3.2	Etat de l'art	99
3.2.1	Approches spatiales	99
3.2.2	Approches fréquentielles	102
3.3	Analyse spectrale du flou	104
3.3.1	Index de qualité	106
3.3.2	Résultats expérimentaux	110
3.4	Analyse perceptuelle	111
3.4.1	Carte de visibilité perceptuelle	111
3.4.2	Index de qualité perceptuel	114
3.4.3	Résultats expérimentaux	116
3.5	Conclusions	117
4	Etude du ringing	119
4.1	Introduction	120
4.2	Etat de l'art	121
4.2.1	Approches spatiales	122
4.2.2	Approches dans le domaine transformé	123
4.3	Estimation du ringing sans référence	125
4.3.1	Analyse locale des gradients	125
4.3.2	Index de qualité	130
4.3.3	Résultats expérimentaux	132

4.4	Conclusion	133
II	Vers un système de mesure universel	137
5	Fusion de métriques	139
5.1	Introduction	140
5.2	Mesure de qualité avec référence	142
5.2.1	Mesures de qualité utilisées	143
5.2.2	Vers une mesure de qualité globale	149
5.2.3	Résultats expérimentaux	153
5.3	Mesure de qualité sans référence multi-dégradations	155
5.3.1	NR-IQM : Mesure proposée	156
5.3.2	Résultats expérimentaux	157
5.4	Conclusions	158
6	Système de mesure complet	161
6.1	Mesure avec référence : Système complet	162
6.1.1	Caractérisation des distorsions	162
6.1.2	Classifieur	164
6.1.3	Résultats expérimentaux	165
6.2	Mesure sans référence : Système complet	168
6.2.1	Caractérisation des distorsions	169
6.2.2	Classifieur	171
6.2.3	Résultats expérimentaux	171
6.3	Conclusions	174
	Conclusion générale	177
	Bibliographie	179

Table des figures

1	Chaîne de communication.	4
1.1	Exemple d'effet de bloc.	13
1.2	Exemple d'effet mosaïque.	14
1.3	Exemple d'effet d'escalier.	14
1.4	Exemple de flou.	15
1.5	Exemple de ringing.	15
1.6	Exemple de faux contours.	16
1.7	Méthode à simple stimulus.	20
1.8	Méthode à double stimulus.	21
1.9	Base IVC : échantillon d'images.	24
1.10	Base LIVE : échantillon d'images.	25
1.11	Base TOYAMA : échantillon d'images.	25
1.12	Base TID 2008 : échantillon d'images	26
1.13	Evaluation de la qualité.	28
1.14	Qualité d'image objective : les différentes approches.	30
1.15	Système Visuel Humain : Schéma simplifié.	36
1.16	Description de l'oeil.	36
1.17	Expérience de Weber-Fechner et tracé de la loi de variation du seuil de perception.	40
1.18	Représentation 2D de l'image fovéale dans le modèle de Moon-Spencer.	41
1.19	Fonction de sensibilité au contraste (CSF).	46
1.20	Transformée en Cortex.	49
1.21	Modèle de masquage de contraste.	51
1.22	Modèle de masquage de contraste.	52
1.23	VDP : Modèle de Daly.	54
1.24	Schéma récapitulatif.	61

2.1	a) Image originale et ses b-c) versions compressées	66
2.2	Schéma synoptique.	71
2.3	Échantillon d'images de la base d'apprentissage.	72
2.4	a) Image originale, b) Image des variances locales, c) Image des gradients nuls et d) Image des gradients pondérés par l'image des variances locales.	73
2.5	Matrice d'accumulation des votes.	75
2.6	Fonction de poids pour les facteurs de qualité de 80, 25, 15 et 10.	76
2.7	a) Image originale b-f) Résultats obtenus pour $F_q=30$, $F_q=20$, $F_q=15$, $F_q=10$ et $F_q=5$, respectivement.	77
2.8	Résultat obtenu sur une image naturelle.	78
2.9	Schéma synoptique.	79
2.10	a) Version binarisée de la carte de prédiction de b) l'image "Beach".	79
2.11	Masque fréquentiel.	80
2.12	a) Image tatouée avec la carte BEP et b) Image tatouée sans la carte BEP.	81
2.13	Distribution du détecteur de sortie avec (droite) et sans (gauche) utilisation de la carte BEP pour l'image "Beach" avec 1000 marques différentes.	82
2.14	Schéma synoptique.	84
2.15	a) Image compressée avec un facteur de qualité égale à 10, b) Régions de la classe C1, c) Régions de la classe C2 et d) Image segmentée.	85
2.16	a) Image compressée et sa b) version segmentée, c) Région R_i de l'image segmentée et d) avec une de ses régions adjacentes N_{ij} , e-j) Cartes de visibilité obtenues pour la R_i et chacune de ses régions voisines N_{ij} , k) Carte de visibilité obtenue pour la région R_i (toutes ses régions adjacentes) et l) Carte de visibilité globale.	86
2.17	Fusion des cartes de visibilité obtenues pour une région R_i donnée.	88
2.18	a) Image synthétique, b) Carte de visibilité et son c) profil 1-D.	89
2.19	Force de filtrage α en fonction de la visibilité.	90
2.20	a) Image compressée JPEG avec une facteur de qualité fixé à 10, b) Image filtrée par la méthode proposée par Zhai et c-d) Images filtrées par la méthode proposée pour la première et troisième itérations, respectivement.	92

2.21 Régions correspondantes aux figures 11.a, b, c et d, respectivement.	92
2.22 Schéma récapitulatif des solutions proposées.	93
3.1 Illustration d'un effet de flou sur un bord net.	96
3.2 Exemple de flou de bougé.	97
3.3 Exemple de flou de mouvement.	97
3.4 Flou dû à une perte d'information.	98
3.5 Flou dû à la compression JPEG.	98
3.6 Un flou dû à un filtrage.	98
3.7 Flou artistique.	99
3.8 a) Image originale et ses versions filtrées b) 1 fois et c) 2 fois par un filtre binomial 3x3.	105
3.9 Spectre d'énergie du a) signal original, b) après un filtrage, c) après deux filtrages successifs.	106
3.10 Schéma synoptique.	107
3.11 a) Image test et b) sa version filtrée, zone agrandie de c) l'image test et d) de sa version filtrée.	107
3.12 a-b) Images avec différents niveaux de flou et c-d) résultats de l'ana- lyse radiale, respectivement.	109
3.13 Index de qualité BI en fonction du DMOS (LIVE).	111
3.14 Schéma synoptique.	112
3.15 a-c) Images tests et ses d-f) versions floutées ainsi que ses g-i) cartes de visibilité correspondantes.	113
3.16 a) Image test, b) carte de visibilité correspondante et sa c) version segmentée, d) Régions affectée à la classe 1 et sa e) version étiquetée, et e) version étiquetée de de l'image c)	115
3.17 a) Image originale et ses b-d) versions dégradées (tache grise).	116
3.18 Index de qualité PBI en fonction du DMOS (LIVE).	117
3.19 Schéma récapitulatif des solutions proposées.	118
4.1 Exemple d'effet de Gibbs 1D sur un signal porte.	120
4.2 Exemple de ringing dans une image.	121
4.3 Détection de contours sur une image compressée par JPEG2000.	124
4.4 Schéma synoptique.	126
4.5 Processus d'obtention de la carte de votes.	127

4.6	a) Image dégradée et sa b) carte de votes correspondante.	127
4.7	Vue paramétrique d'une droite	128
4.8	a) Image dégradée, b) Carte de votes et c) sa table de Hough correspondante.	129
4.9	a) Image des gradients et b) sa version pondérée.	130
4.10	Récapitulatif en images.	131
4.11	a) Image dégradée, b) Carte des gradients pondérés et c) sa version segmentée.	132
4.12	Index de qualité RI en fonction du DMOS.	133
4.13	Index de qualité JNB en fonction du DMOS.	134
4.14	Index de qualité CPBD en fonction du DMOS.	134
4.15	Schéma récapitulatif de la solution proposée.	135
5.1	Dégradations considérées : flou (partie gauche), JPEG (partie centrale) et JPEG2000 (partie droite).	140
5.2	Schéma synoptique.	149
5.3	IQM/D : schéma fonctionnel.	151
5.4	IQM/D : modèle neuronale.	151
5.5	G-IQM : modèle neuronal utilisé.	152
5.6	Schéma de fusion globale	152
5.7	Index de qualité G-IQM en fonction du MOS.	154
5.8	Index de qualité SSIM en fonction du MOS.	155
5.9	Index de qualité pour différentes images : MOS (rouge) et G-IQM (bleu).	155
5.10	Schéma synoptique.	156
5.11	Index de qualité pour différentes images : MOS (rouge) et NR-IQM (bleu).	159
5.12	Schéma récapitulatif.	159
6.1	Schéma synoptique.	163
6.2	Schéma fonctionnel (exemple de distorsion JPEG).	166
6.3	Pourcentage de bonne classification.	166
6.4	Exemple d'erreurs de classification : dégradation 8 confondue avec les artefacts 11 (gauche) et 10 (droite).	168
6.5	Classification de dégradation : Système complet.	168

6.6	Schéma synoptique.	169
6.7	Classifieur utilisé.	172
6.8	Système de mesure sans référence complet.	172
6.9	Pourcentage de bonne classification.	173
6.10	a) Image flou et b) compressée par JPEG2000.	174
6.11	Exemple d'erreur de confusion : classe flou confondue avec la classe JPEG2000.	174
6.12	Système de mesure universel.	175

Liste des tableaux

1.1	Echelle d'évaluation de qualité d'image.	21
1.2	Echelle d'évaluation de qualité d'image à 5 niveaux.	22
1.3	Echelle d'évaluation de la qualité.	22
1.4	Base TOYAMA : conditions de test.	26
1.5	Base TID 2008 : dégradations considérées.	27
1.6	Bases d'images : récapitulatif.	27
2.1	Evaluation de la transparence.	81
2.2	Pourcentage de réussite.	82
2.3	Evaluation de la robustesse (image "beach").	83
3.1	Corrélations obtenues.	110
3.2	Corrélations obtenues.	117
4.1	Résultats de corrélations.	133
5.1	Classement des métriques selon le coefficient de Pearson.	142
5.2	Coefficient de Pearson des métriques VIF et PSNRH.	142
5.3	Métriques de qualité utilisées.	150
5.4	Corrélations de Pearson obtenues pour chaque dégradation.	153
5.5	Corrélations de Pearson obtenues pour <i>G-IQM</i>	154
5.6	Corrélations de Pearson obtenues par chaque métrique (avant fusion).158	
5.7	Corrélations de Pearson obtenues par la méthode proposée, pour chaque dégradation (après fusion).	158
5.8	Corrélations obtenues par la métrique proposée.	158
6.1	Matrice de confusion.	167
6.2	Matrice de confusion.	173

Introduction générale

Nous assistons aujourd'hui à un déploiement formidable des technologies liées aux contenus audio-visuels et en particulier les systèmes d'acquisition et de traitement d'images. La banalisation des outils de manipulation des contenus multimédia a contribué aux mutations socio-économiques et culturelles que connaît le monde d'aujourd'hui. Ces avancées techniques ont en effet changé les pratiques et les exigences humaines. Ainsi, l'observateur des années 60 devant un tube CRT de télévision n'est pas aussi exigeant que son fils des années 2000 devant un écran LCD. Cependant, ces progrès technologiques s'accompagnent le plus souvent de nouvelles problématiques et l'observateur reste toujours insatisfait et en quête d'avantage de qualité et de confort. En effet, plus on progresse au niveau de la résolution et plus on a recours à la compression irréversible qui génère des artefacts propres aux technologies utilisées.

Bien que le développement des systèmes de capture et de restitution d'images ait atteint un niveau de finesse et de rapidité sans précédent, les outils de traitement, de codage et de transmission génèrent diverses dégradations. Elles peuvent être classées en quatre catégories. Il y a d'abord les artefacts inhérents aux systèmes de capture (les bruits d'acquisition et les distorsions optiques) et aux conditions d'acquisition (éclairage, nature de la scène, etc.). Il y a ensuite les artefacts de codage (compression avec perte), les distorsions dues au canal de transmission et enfin les défauts liés au système de décodage et restitution du signal image. En général, seules les distorsions induites par le codeur et le canal de transmission sont prises en compte. La figure 1 résume les différentes phases par lesquelles passe l'image ainsi que les dégradations susceptibles d'apparaître.

Compte tenu de ces limitations, il convient de disposer de systèmes permet-

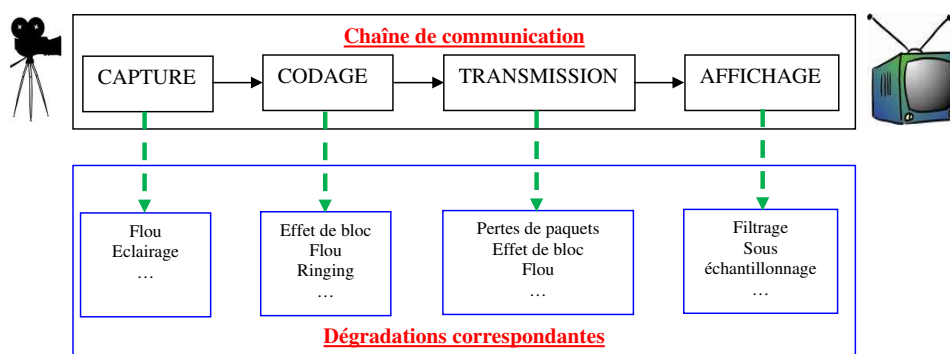


FIGURE 1 – Chaîne de communication.

tant d'évaluer l'impact visuel de ces distorsions sur l'image. Pour répondre à cette demande grandissante, des expériences psycho-visuelles ont été menées pour développer des mesures de qualité. Cette évaluation subjective consiste à mesurer la qualité des images par un panel d'observateurs selon des conditions environnementales bien précises et identiques pour chaque observateur. Ces mesures permettent d'estimer la qualité réelle perçue. Cependant, elles restent coûteuses en temps et difficiles à mettre en place (matériels adéquats, salle,...).

De ce fait, l'évaluation de la qualité objective a été proposée comme solution alternative. Son objectif principal est de remplacer l'évaluation subjective par un outil de mesure automatique, corrélé à l'appréciation visuelle humaine. On distingue alors trois grandes familles de métriques objectives :

Mesures avec référence : Les mesures avec référence sont utilisées pour estimer la qualité entre deux images, l'originale et sa version dégradée. Généralement, aucune information liée au type de dégradation n'est exploitée. Elles sont donc supposées universelles. La majorité des métriques proposées dans la littérature sont avec référence.

Mesures sans référence : Orientées pour une dégradation spécifique, les mesures sans référence sont les plus attrayantes d'un point de vue applicatif. En effet, l'image dégradée est la seule donnée nécessaire. Cependant, le type de distorsion est en général supposé connu.

Mesures avec référence réduite : Les mesures avec référence réduite ont été développées afin de pallier les contraintes liées aux métriques avec et sans référence. Ce type de mesure nécessite uniquement certaines caractéristiques de l'image originale. De plus, aucune information à priori n'est généralement exploitée. Elles restent les méthodes les moins développées.

Ainsi, l'application visée déterminera le type de métrique à utiliser.

0.1 Contexte

Dans le cadre de cette thèse, nous nous sommes intéressés tout d'abord aux trois dégradations les plus répandues, à savoir les effets de bloc, le flou et le ringing (voir chapitres 2, 3 et 4 respectivement). Ce travail s'inscrit dans le projet HD3D.iio de Cap Digital [htt]^[1] dans lequel participent des partenaires industriels et académiques, l'aspect qualité d'image prend ainsi tout son sens dans le domaine du cinéma. En effet, lors de la commercialisation d'un film numérique sous format DVD, une étape de compression est indispensable. Cette compression engendre, selon le taux de compression, des dégradations assez gênantes pour l'observateur. Elle est aussi appliquée lors des échanges entre les différents partenaires pour une première visualisation notamment.

La DCI [Ini08]^[2], qui préconise certaines recommandations de création, de transport, de projection et de protection des films dédiés au cinéma numérique, propose d'utiliser JPEG2000 comme standard de compression (image par image). Ce standard de compression engendre des dégradations telles que le ringing et le flou. Le contexte de cette thèse répond ainsi à des objectifs à la fois industriels et académiques.

Nous inscrivons notre travail dans une démarche où l'on tient compte à la fois de la nature du signal image et de l'observateur. Nous pensons aussi comme plusieurs

[1] <http://www.hd3d.fr/en/>.

[2] Digital Cinema Initiatives. Digital Cinema System Specification. *LLC, Member Representatives Committee*, 2008.

auteurs et notamment [Kle93]^[1][JAC93]^[2] que la qualité d'image est un problème multidimensionnel et qu'il convient de l'appréhender en intégrant plusieurs aspects liés à la dégradation observable et à l'outil de mesure. Nous pensons qu'il est illusoire de tenter de quantifier toutes les dégradations au moyen d'un seul index de qualité comme le prétendent implicitement plusieurs méthodes connues. C'est donc dans le cadre de cette réflexion que nous développons cette étude.

Ainsi, nous proposons de combiner de façon coopérative les métriques connues de manière à converger vers un système plus ou moins universel. Nous intégrons dans cette approche l'aspect apprentissage très lié à la notion de qualité. En effet, notre faculté à reconnaître le flou, le bruit et d'autres artefacts dans une image dégradée sans avoir recours à sa version originale, montre bien que ce capital de connaissance est acquis lors du processus d'apprentissage.

Dans notre démarche, nous abordons alors le problème de qualité en le décomposant en trois étapes toutes décisives :

- Caractérisation des différentes dégradations.
- Prédiction et estimation des distorsions au moyen de métriques dédiées (les plus adéquates).
- Combinaison des métriques selon un schéma coopératif via un processus d'apprentissage.

Ces trois phases constituent les axes principaux de cette recherche.

0.2 Organisation de la thèse

Ce manuscrit est composé de 6 chapitres.

Chapitre 1 : État de l'art L'objectif dans ce chapitre est de définir les notions essentielles liées à la qualité d'image. L'évaluation subjective y est

-
- [1] S.A. Klein . Image quality and image compression: a psychophysicist's viewpoint. *Chapter in Digital Images and Human Vision*, ed. AB. Watson, PIT Press, pp. 73-88, 1993.
- [2] A. J Ahumada, Jr. and C. H. Null. Image quality: A multidimensional problem. *IEEE Signal Processing Magazine*, Vol. 18, pp. 851-854, 1992.

tout d'abord présentée à travers les différents protocoles de mesure de qualité [ITU96] [VQE00]^[3]. Un ensemble de métriques de qualité est ensuite présenté ainsi qu'une description succincte des propriétés les plus pertinentes du Système Visuel Humain (SVH).

Chapitre 2 : Effets de blocs Ce chapitre est consacré à l'étude des dégradations produites par les méthodes de compression par bloc. Nous présentons dans un premier temps les origines de cette dégradation ainsi qu'un bref état de l'art. Les méthodes de prédiction (avant compression) et de réduction proposées y sont décrites et discutées en détails.

Chapitre 3 : Le Flou Après avoir présenté les origines de cette dégradation, un bref état de l'art des méthodes d'estimation du flou est présenté en première partie de ce chapitre. Une méthode d'estimation du flou basée sur une analyse radiale du spectre de fréquence y est exposée. Une seconde métrique basée sur l'exploitation d'attributs visuels du SVH est aussi décrite.

Chapitre 4 : Le Ringing Ce chapitre est dédié à l'analyse de l'effet de Gibbs 2D ou "ringing" qui résulte notamment d'une compression par JPEG2000 [TM01]^[4]. Comme pour les deux chapitres précédents, après avoir décrit les causes d'apparition de la dégradation et présenté un bref état de l'art, une mesure basée sur la combinaison de plusieurs cartes (gradient, Hough, ...) y est décrite en détails.

Chapitre 5 : Fusion des métriques Dans ce chapitre, notre étude s'est focalisée sur l'amélioration des performances des métriques de qualité existantes. La première partie est consacrée aux mesures avec référence où un nouveau schéma de fusion permettant d'estimer la qualité de manière plus efficace est présenté. La deuxième partie, quant à elle, est dédiée aux mesures sans référence. Une métrique multi-dégradations sans référence est proposée.

[3] VQEG Final report from the video quality experts group on the validation of objective models of video quality assessment. <ftp://vqeg.its.bldrdoc.gov/Documents/>, 2000.

[4] D. Taubman and M. Marcellin. JPEG2000: Image compression fundamentals, standards and practice. *Kluwer Academic Publishers*, 2001.

Chapitre 6 : Système de mesure complet Dans ce chapitre, les résultats obtenus précédemment (chapitre 5) sont repris et améliorés. Une phase de détection du type de dégradation contenu dans l'image est introduite. Cette étape permet de sélectionner la métrique la plus appropriée et *in fine* de mieux estimer la qualité. Notre étude s'est limitée ici aux mesures avec et sans référence.

Une conclusion générale ainsi que des perspectives sont données à la fin de ce manuscrit.

0.3 Liste des publications et communications

Ce travail de thèse a donné lieu aux publications suivantes :

Article journal

1. **Aladine Chetouani**, Ghiles Mostafaoui, Azeddine Beghdadi,
Deblocking filtering method using a perceptual map
Elsevier Journal Signal Processing : Image Communication - Special Issue on Image and Video Quality Assessment, Vol. 25, Issue 7, pp. 527-534, 2010.

Conférences internationales avec actes

1. **Aladine Chetouani**, Azeddine Beghdadi
Image quality assessment based on distortion identification
SPIE - Electronic Imaging - Image Quality and System Performance VIII, 2011.
2. **Aladine Chetouani**, Azeddine Beghdadi, Mohamed Deriche
A universal full reference image quality metric based on a neural fusion approach
IEEE-ICIP, pp. 2517-2520, 2010.
3. **Aladine Chetouani**, Azeddine Beghdadi, Mohamed Deriche
Classification of image degradation using multiple image quality metrics and

linear discriminant analysis

EUSIPCO, 2010.

4. **Aladine Chetouani**, Azeddine Beghdadi, Mohamed Deriche
Statistical modeling of image degradation based on quality metrics
IEEE-ICPR, pp. 714-717, 2010.
5. **Aladine Chetouani**, Azeddine Beghdadi, Mohamed Deriche
Image distortion analysis and classification scheme using a neural approach
IEEE-EUVIP, pp. 183-186, 2010. **Best Student Paper Award**.
6. **Aladine Chetouani**, Phi-Bang Nguyen, Marie Luong and Ghiles Mostafaoui
Content-based watermarking robust to low bit rate jpeg compression
IEEE-ISSPA, pp. 29-32, 2010.
7. Mohamed Deriche, Azeddine Beghdadi, **Aladine Chetouani**
A distortion-based ranking of image quality indices using mutual information
IEEE-ISSPA, pp. 484-487, 2010.
8. **Aladine Chetouani**, Azeddine Beghdadi, Shaohua Chen, Ghilès Mostafaoui
A free reference image quality measure using neural networks
VPQM 2010.
9. Shaohua Chen, Azeddine Beghdadi, **Aladine Chetouani**
A new color image quality index
VPQM 2010.
10. **Aladine Chetouani**, Ghiles Mostafaoui, Azeddine Beghdadi
Deblocking method using a perceptual recursive filter
IEEE-ICIP, pp. 3925-3928, 2009.
11. **Aladine Chetouani**, Ghiles Mostafaoui, Azeddine Beghdadi
A new free reference image quality index based on perceptual blur estimation
IEEE-PCM, pp. 1185-1196, 2009. **Published in Lecture notes in Computer Science Series, Vol. 5876 (Springer)**.
12. **Aladine Chetouani**, Azeddine Beghdadi, Mohamed Deriche
A new no reference image quality index for blur estimation in the frequency domain
IEEE-ISSPIT, pp. 155-159, 2009.

13. **Aladine Chetouani**, Ghiles Mostafaoui, Azeddine Beghdadi
Predicting blocking effect in the spatial domain using a learning approach
IEEE-SIGMAP, pp. 197-201, 2008.
14. Shaohua Chen, Azeddine Beghdadi, **Aladine Chetouani**
Multi-scale and scalable video quality assessment
IEEE-ICCE, pp. 1-2, 2008.

Conférences internationales sans actes

1. **Aladine Chetouani**, Azeddine Beghdadi, Ghiles Mostafaoui, Mohamed Deriche
Mesure de flou sans référence dans le domaine spectrale
ISSSPA/WSIP'09.
2. Zehira Haddad, Azeddine Beghdadi, **Aladine Chetouani**, Amina Serir, Anissa Mokraoui
Nouvelle mesure de qualité d'image basée sur les curvelets
ISSSPA/WSIP 2009.
3. **Aladine Chetouani**, Ghiles mostafaoui, Azeddine Beghdadi
Analyse et Prédiction des artefacts de compression
JDTIC 2009.

Conférence nationale avec actes

1. **Aladine Chetouani**, Ghiles mostafaoui, Azeddine Beghdadi
Méthode de prédiction des effets de blocs dans les standards de compression
CORESA 2009.

Rapports techniques et livrables

1. État de l'art sur la qualité d'image.
2. Interface graphique de mesure de qualité objective (langage C).
3. Tests subjectifs : Impact de la compression JPEG2000 sur les images HD.

Chapitre 1

La qualité d'image

Sommaire

1.1	Introduction	13
1.2	Notions et définitions	17
1.3	Evaluation subjective	18
1.3.1	Les facteurs influant	18
1.3.2	Les protocoles	20
1.3.2.1	Simple stimulus	20
1.3.2.2	Double stimuli	21
1.3.2.3	Comparatif	22
1.3.3	Bases d'images dédiées à l'évaluation de performance . .	24
1.3.3.1	IVC	24
1.3.3.2	LIVE	24
1.3.3.3	Toyama	25
1.3.3.4	TID 2008	26
1.4	Evaluation objective	28
1.4.1	Approches basées EQM	30
1.4.2	Approches structurelles	33
1.4.3	Approches inspirées du Système Visuel Humain	35
1.4.3.1	Le Système Visuel humain	35
1.4.3.2	Les modèles	38
1.4.3.3	Métriques inspirées du SVH	54

1.4.4	Approches neuronales	58
1.5	Conclusion	59

1.1 Introduction

Tout au long de la chaîne de communication, le signal image subit différents traitements nécessaires qui permettent une production et une diffusion plus aisées. Cependant, ces traitements ne garantissent pas toujours la restitution exacte du signal et induisent malheureusement des dégradations qui impactent sur la qualité visuelle de l'image. Dans ce qui suit, nous présentons en premier lieu quelques distorsions connues pour ensuite introduire les notions liées à la qualité d'image. La liste présentée ci-dessous n'étant pas exhaustive, nous invitons les lecteurs à se référer à [Win05]^[1] [WR05]^[2] pour plus de détails.

Effet de bloc : Les effets de bloc font référence à l'apparition de bloc dans l'image.

Cette dégradation est due à une quantification indépendante des blocs conduisant ainsi à des discontinuités au voisinage de ces blocs. La distorsion se traduit visuellement par l'apparition d'un effet de mosaïque. La régularité de cet artefact la rend très gênante (cette dégradation est étudiée plus en détails dans le chapitre 2). Un exemple est présenté par la figure 2.1



FIGURE 1.1 – Exemple d'effet de bloc.

Effet mosaïque : L'origine de cette dégradation est la perte partielle ou totale du contenu d'un bloc de l'image due à la compression bas débit par exemple,

-
- [1] S. Winkler. Digital Video Quality : Vision Models and Metric. *John Wiley and Sons*, 2005.
 - [2] H.R. Wu and K.R. Rao. Digital Video Image Quality and Perceptual Coding. *CRC Press*, 2005.

ou encore au faible débit du canal de transmission. Cet effet de mosaïque est plus visible dans les zones texturées.



FIGURE 1.2 – Exemple d'effet mosaïque.

Effet d'escalier ("staircase effect") : Due à l'approximation des contours diagonaux par des contours horizontaux et verticaux, cette dégradation se traduit visuellement par la formation d' "escalier " comme illustré par la figure 1.3. Elle est due essentiellement à une troncation élevée des coefficients hautes fréquences. Cet effet est bien illustré par l'exemple de la figure 1.3.



FIGURE 1.3 – Exemple d'effet d'escalier.

Le flou : Cet artefact se manifeste par une perte de finesse et de visibilité des détails résultant d'une diminution du contraste, comme illustré par la figure 1.4. En compression, cela est dû à l'étape de sélection des coefficients et à une quantification brutale. Au niveau des composantes de chrominances, le flou se manifeste par un étalement de la couleur ("color bleeding"). Cet artefact

sera analysé un peu plus en détails dans le chapitre 3.



FIGURE 1.4 – Exemple de flou.

Fausses couleurs : Le traitement des composantes de luminance et de chrominances de façon indépendante est à l'origine de cette distorsion. En effet, les images couleurs sont généralement compressées composante par composante. De ce fait, la quantification d'une composante indépendamment des autres peut induire de fausses couleurs.

Effet de "ringing" : Communément appelé effet de Gibbs, le "ringing" est associé à l'apparition d'oscillations à proximité des régions à fort contraste. Cet artefact est la conséquence directe de la quantification des coefficients hautes fréquences. Sa visibilité est accrue le long des contours à fort contraste proches des zones homogènes et dépend fortement de la valeur du contraste. Cette dégradation est étudiée en détails dans le chapitre 4.



FIGURE 1.5 – Exemple de ringing.

Faux contours (false contouring) : Cette distorsion fait émerger de faux contours dans l'image. Elle se manifeste notamment lors d'une forte compression par bloc de l'image. Sa visibilité est amplifiée dans les zones homogènes. Cet artefact apparaît aussi dans les régions d'activité spatiale élevée mais reste moins gênante (effet de masquage). La figure 1.6 présente un exemple d'image contenant cette distorsion.



FIGURE 1.6 – Exemple de faux contours.

Effet moustique : L'effet de moustique est un artefact temporel visible principalement dans les régions à fort contraste. Il se manifeste sous forme de fluctuations du signal au voisinage des contours. Cette distorsion est liée à une mauvaise estimation de mouvement ou à une quantification brutale des coefficients.

Cette présentation sommaire permet de rendre compte de la nécessité de disposer d'outils efficaces permettant de mesurer l'impact des distorsions sur la qualité visuelle de l'image. Différentes études ont été réalisées dans ce sens, essayant ainsi d'apporter des solutions à cette problématique. Dans la suite de ce chapitre, nous décrivons les solutions proposées en définissant d'abord quelques notions importantes.

1.2 Notions et définitions

Avant d'aborder l'état de l'art sur les méthodes d'évaluation de la qualité d'image, rappelons d'abord quelques notions importantes pour lever toute ambiguïté. Le plus souvent on confond par exemple fidélité et qualité.

La fidélité : La fidélité se réfère à la notion de conservation de l'information. Il s'agit de mesurer la similarité entre un stimulus original et une version à évaluer.

La qualité : La qualité d'image est quant à elle une notion très liée à la perception visuelle humaine. Il y a aussi une dimension liée à notre faculté d'apprentissage. En effet, notre capacité à évaluer la qualité d'une image sans avoir recours à l'originale tient de ce que notre système possède la faculté d'emmagasiner et d'apprendre une quantité énorme d'information. Évaluer la qualité d'une image c'est lui associer un ou plusieurs qualificatifs permettant de situer sa position relative dans un référentiel défini par notre sens et selon l'application envisagée. Ainsi une image de mammographie contenant des microcalcifications, qui apparaissent comme des points aberrants, peut être jugée de mauvaise qualité par un pur traiteur du signal. Alors qu'un radiologiste, au contraire, interprète les microcalcifications comme étant une information pertinente. La qualité d'image n'est pas seulement liée à l'apparence visuelle mais dépend aussi de l'application et de l'usage. Dans le cadre de cette thèse nous nous intéressons uniquement à la qualité perceptuelle. Il existe essentiellement deux types d'évaluation.

Evaluation subjective : L'évaluation subjective fait appel à des observateurs humains (utilisateurs finaux) pour évaluer (ou comparer) la qualité d'une image (ou de plusieurs images) selon un protocole bien défini. Elle est considérée comme le moyen le plus fiable et donc la référence pour comparer les différentes métriques proposées.

Evaluation objective : L'évaluation objective fait référence aux méthodes basées sur l'analyse et la mesure quantitative du niveau de dégradation au

moyen de métriques directement liées au signal physique. La tendance actuelle est d'exploiter au mieux les connaissances des mécanismes de la perception visuelle de façon à tendre vers les résultats de l'appréciation subjective. Les métriques de qualité objectives sont alors évaluées à travers l'étude de la cohérence avec les notes subjectives.

En s'appuyant sur ces définitions, nous présentons dans ce qui suit les méthodes d'évaluations de qualité subjectives et objectives.

1.3 Evaluation subjective

Considérée comme le moyen le plus fiable pour mesurer la qualité réelle, l'évaluation subjective fait appel directement à l'observateur humain. Ce dernier est amené à juger la qualité des images qui lui sont présentées selon une grille d'évaluation à plusieurs niveaux d'appréciation. A l'issue de ces tests, une note subjective appelée MOS (Mean Opinion Score) est obtenue.

Lors de la réalisation de ces tests, il est important de porter une attention particulière à certains facteurs pouvant altérer les jugements des participants. Ces facteurs sont énumérés dans ce qui suit.

1.3.1 Les facteurs influant

Pour s'assurer de la fiabilité de l'évaluation subjective, il convient de tenir compte de certains facteurs pouvant influencer grandement le jugement des observateurs [Bek99]^[1]. Nous listons ici les plus pertinents.

Distance d'observation : La visibilité d'un stimulus dépend fortement de la distance d'observation. Il est recommandé de fixer cette distance entre 4 à 6 fois la hauteur de l'image de projection et de la maintenir tout au long du test.

[1] N. Bekkat. Critère objectifs de qualité subjective d'images monochromes. Conception du modèle et validation expérimentale. *Rapport de thèse, Université de Nantes - IRESTE*, 1999.

Conditions de visualisation : L'éclairage de la salle ainsi que l'environnement (couleur de fond) jouent également un rôle important dans la perception des stimuli. Un éclairage élevé risque d'éblouir l'observateur et ainsi altérer sa perception. De même, la couleur de fond de la salle a tendance à affecter l'apparence de certaines couleurs.

Écran : Il est nécessaire de calibrer l'écran de façon à reproduire les couleurs souhaitées. Sans cette calibration, les couleurs affichées peuvent être différentes du stimulus d'origine. Il est important aussi de tenir compte de quelques artefacts liés à la technologie de l'écran utilisé [Tou09]^[2].

Le choix des images : Un panel d'images de contenus visuels assez variés et le plus représentatif possible des scènes liées à l'application visée doit être utilisé.

Facteurs psychologiques : Pour minimiser les effets liés à l'état psychologique de l'observateur, il convient de débiter les tests par une séance d'initiation permettant à l'observateur de mieux appréhender le test. Il est aussi important d'expliquer clairement l'objectif du test et le protocole à respecter sans pour autant influencer le jugement des participants. De plus, il est préférable de mener de nombreux tests de courte durée ($\sim 10min$) et variés plutôt que de contraindre l'observateur à des séances longues ($\sim 30min$) et fatigantes qui risquent de l'amener à un jugement rapide et aberrant.

Les observateurs : Le choix des observateurs doit être établi en fonction de l'application visée par les tests et le plus représentatif possible de la population moyenne concernée. Des tests d'acuité visuelle doivent être réalisés sur l'ensemble des observateurs. Un nombre suffisant de participants doivent participer aux tests. Il est généralement admis que les tests sont statistiquement valides si ce nombre est supérieur ou égal à 16.

[2] S. Tourancheau. Caractérisation objective et modélisation psychovisuelle du flou de mouvement sur les écrans à cristaux liquides. Impact sur la qualité perçue. *Rapport de thèse, Université de Nantes*, 2009.

Ainsi, pour minimiser les erreurs de jugements, tous ces paramètres doivent être pris en considération et notés soigneusement. Notons que l'évaluation subjective de la qualité d'image dépend d'autres facteurs encore mal maîtrisés tel que l'humeur, l'âge, la culture, le niveau intellectuel. Ici, nous nous sommes limités à quelques facteurs seulement.

1.3.2 Les protocoles

Différentes méthodes d'évaluation subjective ont été définies par l'ITU [Dur]^[1]. On recense essentiellement trois grandes familles : les tests à simple stimulus permettant de juger de la qualité d'image sans aucune information relative à l'image originale. Les tests à double stimuli où il est demandé à l'observateur de juger la similarité entre deux images et les tests comparatifs dont l'objectif principal est de comparer deux ou plusieurs stimuli. Ces méthodes sont décrites ci-après.

1.3.2.1 Simple stimulus

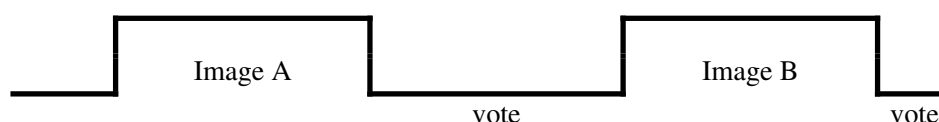


FIGURE 1.7 – Méthode à simple stimulus.

La méthode à simple stimulus, appelé "Single Stimulus Continuous Quality Scale (SSCQS)" permet de juger de la qualité d'un stimulus à la fois. Il est à noter que le mode de présentation, illustré par la figure 1.7, se voit ainsi limité. Les images sont présentées une par une avec un temps de latence entre deux présentations. Ce temps permet à l'observateur de noter la qualité de l'image.

[1] Durant. Méthodes d'évaluation subjective de la qualité audiovisuelles pour applications multimédias. *Recommandation UIT-T P.911*.

Excellente
Bonne
Assez bonne
Médiocre
Mauvaise

TABLE 1.1 – Echelle d'évaluation de qualité d'image.

Des échelles à 5, 6, 7 ou 100 niveaux d'appréciation peuvent être utilisées. Le tableau 1.1 présente un exemple d'échelle.

1.3.2.2 Double stimuli

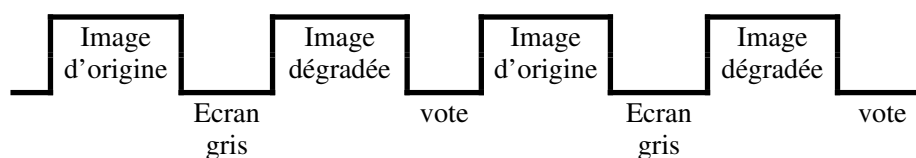


FIGURE 1.8 – Méthode à double stimulus.

Contrairement à la méthode à simple stimulus, cette méthode, également appelée "Double Stimuli Continuous Quality Scale (DSCQS)", a pour objectif de mesurer la qualité (ou fidélité perceptuelle) d'un stimulus par rapport à sa version originale. L'image de référence est d'abord présentée, suivi d'un écran gris et de sa version dégradée (ou pas). Un second écran gris est affiché, il permet à l'observateur de disposer d'un temps de vote. Un exemple de présentation est illustré par la figure 1.8. Il est à noter que la durée de présentation doit être identique pour les deux stimuli. Le mode de présentation peut aussi être modifié (originale/dégradée, dégradée/originale, dégradée/dégradée et originale/originale).

Différentes échelles d'appréciation peuvent être utilisées où chaque note désigne la qualité ou le niveau de perception de la dégradation. Un exemple d'échelle de mesure à 5 niveaux est illustré par le tableau 1.2.

Notes	Qualité	Dégradation
5	Excellente	Imperceptible
4	Bonne	Perceptible
3	Passable	Peu dégradée
2	Médiocre	Dégradée
1	Mauvaise	Très dégradée

TABLE 1.2 – Echelle d'évaluation de qualité d'image à 5 niveaux.

Beaucoup moins bon
Moins bon
Identique
Légèrement mieux
Mieux
Beaucoup mieux

TABLE 1.3 – Echelle d'évaluation de la qualité.

1.3.2.3 Comparatif

L'objectif des méthodes comparatives est de comparer deux ou plusieurs images selon leur qualité. La différence majeure réside dans le fait que l'on souhaite ici quantifier la perception des différences entre des stimuli (voir tableau 1.3). Ainsi, aucune information relative à la qualité des stimuli n'est disponible. De ce fait, l'image jugée la meilleure peut être de mauvaise qualité.

Il existe aussi des méthodes dédiées à l'évaluation subjective des vidéos (peuvent être utilisées aussi pour l'image). On peut citer par exemple la méthode SAMVIQ "Subjective Assessment Methodology for Video Quality" [Bli03]^[1]. Cette méthode a été développée par France Télécom R&D et standardisée par l'Union Européenne de radio télévision. Elle offre la possibilité à l'observateur de visionner les séquences dans l'ordre qu'il le souhaite et suivre son propre rythme. La modification des notes et la répétition des séquences ainsi que le nombre de visionnages ne sont pas limités.

Une fois les tests réalisés, les résultats sont analysés et combinés de manière à obtenir une note unique par image décrivant sa qualité moyenne. Cette note appelée "Mean Opinion Score (MOS)" est donnée par la formule suivante :

[1] J.L. Blin. Samviq - Subjective assessment methodology for video quality. *rapport technique BPN 056, EBU Project Group B/VIM Video Multimedia*, 2003.

$$MOS(i) = \frac{1}{N_{obs}} \sum_{j=1}^{N_{obs}} Note_i(j) \quad (1.1)$$

où N_{obs} est le nombre total de participants et $Note_i(j)$ la note affectée à l'image i par l'observateur j .

Un intervalle de confiance est souvent associé à chaque note MOS, permettant ainsi de réduire l'impact d'éventuelles erreurs. Il est généralement fixé à 95% et est donné par la relation suivante :

$$[MOS(i) - e_j, MOS(i) + e_j] \quad (1.2)$$

où

$$e_j = 1.96 \cdot \sigma_j \quad (1.3)$$

et

$$\sigma_j = \sqrt{\frac{1}{N_{obs} - 1} \cdot \sum_{w=1}^{N_{obs}} (Note_i(k) - MOS(k))^2} \quad (1.4)$$

Une analyse plus approfondie permet de détecter et de rejeter d'éventuelles erreurs. On peut par exemple vérifier la cohérence de l'appréciation subjective [CLS05]^[2] [BT.02]^[3]. Il suffit pour cela de comparer les notes affectées à une même image.

-
- [2] C. Charrier, C. Larabi and H. Saadane. Evaluation de la qualité des images. *Ecole d'hiver sur l'image numérique couleur*, 2005.
 - [3] BT.500-11 Methodology for the subjective assessment of the quality of television pictures. *International Telecommunication Union*, 2002.

1.3.3 Bases d'images dédiées à l'évaluation de performance

Depuis quelques années, plusieurs bases d'images avec des tests d'évaluation subjective associés ont été proposées. Pour chacune des bases ci-dessous, nous décrivons le type d'images, les dégradations considérées, le protocole choisi et les données mises à disposition des utilisateurs (MOS, notes par observateur, ...).

1.3.3.1 IVC

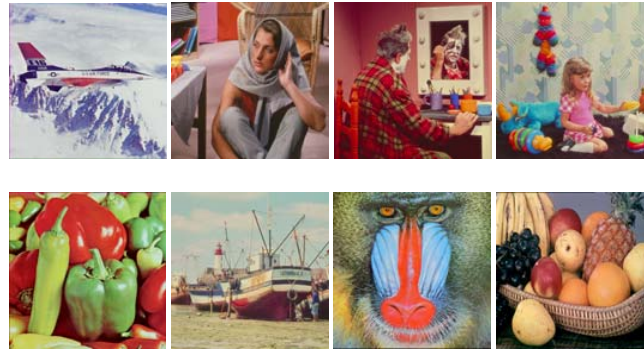


FIGURE 1.9 – Base IVC : échantillon d'images.

La base IVC [CA05]^[1] contient 10 images de référence ainsi que leurs versions dégradées (voir figure 1.9). Le nombre de dégradations considérées est ici de 4 : JPEG, JPEG2000, "LAR coding" et le flou. Ainsi, 235 images ont été jugées par 15 observateurs. La distance d'observation a été fixée à 6 fois la hauteur de l'image. Le protocole utilisé est à double stimuli avec une échelle à 5 niveaux.

1.3.3.2 LIVE

Proposée dans [SWCB06]^[2], cette base contient 29 images de référence ainsi que les versions dégradées. Le nombre de distorsions considérées est ici de 5 :

-
- [1] P. Le Callet and F. Autrusseau. Subjective quality assessment IRCCyN/IVC database. <http://www.irccyn.ec-nantes.fr/ivcdb/>, 2004.
 - [2] H.R. Sheikh, Z. Wang, L. Cormack and A.C. Bovik. LIVE Image Quality Assessment Database. <http://live.ece.utexas.edu/research/quality>, 2006.

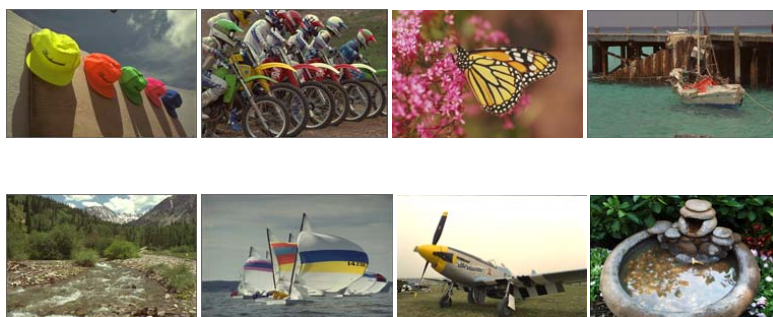


FIGURE 1.10 – Base LIVE : échantillon d’images.

JPEG2000, JPEG, Bruit blanc, Flou et erreurs de transmission dans JPEG2000. Un total de 982 images dont 203 images de référence ont été jugées par une vingtaine d’observateurs. Un échantillon d’images de cette base est présenté figure 1.10. Le protocole SSCQS a été adopté dans ces tests. Une échelle continue linéaire divisée en 5 niveaux a été utilisée. Les notes brutes de chaque observateur ont été converties en scores de différence (DMOS : Difference Mean Opinion Score). La force de la dégradation ajoutée (taux de compression, écart-type du bruit,...) est aussi précisée.

1.3.3.3 Toyama

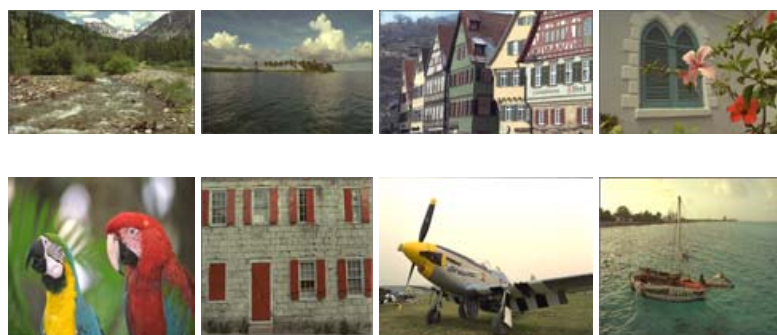


FIGURE 1.11 – Base TOYAMA : échantillon d’images.

Ici, l'expérience psycho-visuelle avait pour objet 2 dégradations : JPEG et JPEG2000 [SkH08]^[1]. Un panel de 16 observateurs non-experts a été invité à attribuer à chaque image une note comprise entre 1 et 5 (respectivement la plus mauvaise et la meilleure qualité). Le nombre d'images par dégradation s'élève à 98. Un échantillon d'images de référence est présenté par la figure 1.11. Deux versions sont disponibles, une première version avec écran CRT (Japon) et une version avec écran LCD (France). Les conditions de test de la version japonaise sont répertoriées dans le tableau ci-dessous,

Méthode	ACR
Échelles d'évaluation	5 niveaux
Nombre d'images de référence	14
Dégradation considérées	JPEG et JPEG2000
Nombre d'observateurs	16 (non expert)
Distance d'observation	4H (H : hauteur de l'image)

TABLE 1.4 – Base TOYAMA : conditions de test.

1.3.3.4 TID 2008



FIGURE 1.12 – Base TID 2008 : échantillon d'images

Dans [PLZ⁺09]^[2], des tests subjectifs ont été menés dans différents pays : Finlande (251 expériences), Italie (150 expériences) et Ukraine (437 expériences). Au

[1] Z.M. Parvez, Y. kawayoke and Y. Horita. Image Quality Evaluation Database. *mict toyama university*, 2008.

[2] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli and F. Battisti. TID2008 -

total, 1700 images dégradées sont dérivées de 25 images de référence. Le nombre de dégradations considérées est de 17, soit 100 images par dégradation (voir tableau 1.5).

Dégradation	Type de dégradation
1	Additive Gaussian noise
2	Additive noise in color components
3	Spatially correlated noise
4	Masked noise
5	High frequency noise
6	Impulse noise
7	Quantization noise
8	Gaussian blur
9	Image denoising
10	JPEG compression
11	JPEG2000 compression
12	JPEG transmission errors
13	JPEG2000 transmission errors
14	Non eccentricity pattern noise
15	Local block-wise distortions of different intensity
16	Mean shift (intensity shift)
17	Contrast change

TABLE 1.5 – Base TID 2008 : dégradations considérées.

Une échelle à 10 niveaux a été utilisée où les notes 0 et 9 correspondent, respectivement, à la qualité la plus mauvaise et la meilleure.

Le tableau 1.6 présente un récapitulatif des caractéristiques les plus importantes des bases d'images décrites précédemment.

<i>Base</i>	<i>Protocole</i>	<i>Nb de dég.</i>	<i>Dist d'obser.</i>	<i>Nb d'obser.</i>	<i>Echelle</i>
IVC	DSIS	4	6H	15	1-5
LIVE	SSCQS	5	–	20-29	0-10
TOYAMA	ACR	2	4H	16	1-5
TID 2008	–	17	–	838	0-9

TABLE 1.6 – Bases d'images : récapitulatif.

A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics. *Advances of Modern Radioelectronics, Vol. 10, pp. 30-45, 2009.*

Dans le cadre de cette thèse, nous avons utilisé essentiellement trois bases : LIVE, IVC, TID 2008. Ces bases nous semblent les plus appropriées et les plus riches à l'heure actuelle. Il est à noter que d'autres bases d'images sont aussi disponibles [CH07a]^[1] [Cha10]^[2].

Tous ces tests nécessitent un matériel et des outils informatiques spécifiques (salle dédiée, environnement, éclairage, écran, sonde de calibration, observateurs,...) rendant ainsi l'évaluation subjective lourde, complexe et difficilement reproductible. Toutes ces raisons ont conduit à l'élaboration de mesures objectives permettant d'évaluer automatiquement la qualité des images.

1.4 Evaluation objective

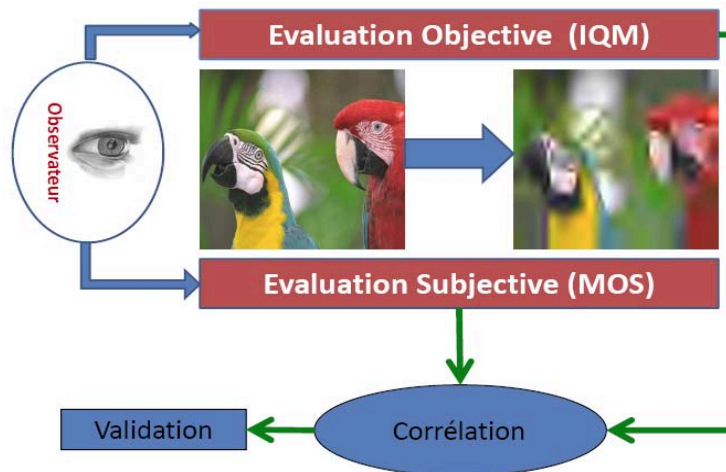


FIGURE 1.13 – Evaluation de la qualité.

Bien que l'évaluation subjective reste la mesure de référence, il paraît indispensable de disposer d'outils d'estimation de la qualité d'image moins contraignants. Plusieurs travaux ont été menés dans ce sens, donnant ainsi naissance à

-
- [1] D. Chandler and S. Hemami. Subjective image database.
<http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html>.
 - [2] D. Chandler. Content-Based Strategies of Image and Video Quality Assessment.
<http://vision.okstate.edu/index.php?loc=csiq>, 2010.

une pléthore de méthodes d'évaluation de la qualité d'image objective. Certaines sont basées sur des approches purement mathématiques telles que l'erreur quadratique moyenne ou encore des distances de type Kullback-Leibler. D'autres intègrent quelques caractéristiques et descripteurs de l'image (le contraste, le contenu fréquentiel, les structures géométriques,...) ou tentent de reproduire les processus d'appréciation de la qualité d'image via des modèles des mécanismes du SVH.

Les mesures de qualité objective proposées dans la littérature peuvent être classées en 3 catégories selon que l'on dispose ou non de l'image originale. On distingue alors les métriques avec référence, sans référence et avec référence réduite.

Mesures avec référence : Ce type de mesure permet de mesurer la qualité entre une image originale et sa version dégradée. Généralement, aucune information liée au type de dégradation n'est exploitée. Elles sont donc supposées être universelles. La majorité des métriques développées sont avec référence.

Mesures sans référence : Orientées pour une dégradation spécifique, les mesures sans référence sont les plus attrayantes car elles permettent d'estimer la qualité uniquement à partir de l'image dégradée. Cependant, le type de dégradation contenu dans l'image est généralement supposé connu.

Mesures avec référence réduite : Les mesures avec référence réduite ont été développées afin de pallier les contraintes liées aux mesures avec et sans référence, en offrant une solution intermédiaire. Ce type de mesure permet de quantifier la qualité d'une image à partir de quelques attributs (ou descripteurs) de l'image originale et sans aucune information a priori sur le type de dégradation contenu dans l'image. Bien que dans un cadre applicatif ces méthodes soient relativement accessibles, elles restent les moins prisées.

Le type de métriques de qualité d'image dépend de l'application visée. Dans le cadre de cette thèse, nous nous sommes intéressés aux mesures avec et sans référence.

Nous proposons de classer les métriques en quatre sous-ensembles : Les métriques basées sur l'erreur quadratique moyenne (EQM), celles basées sur l'analyse des

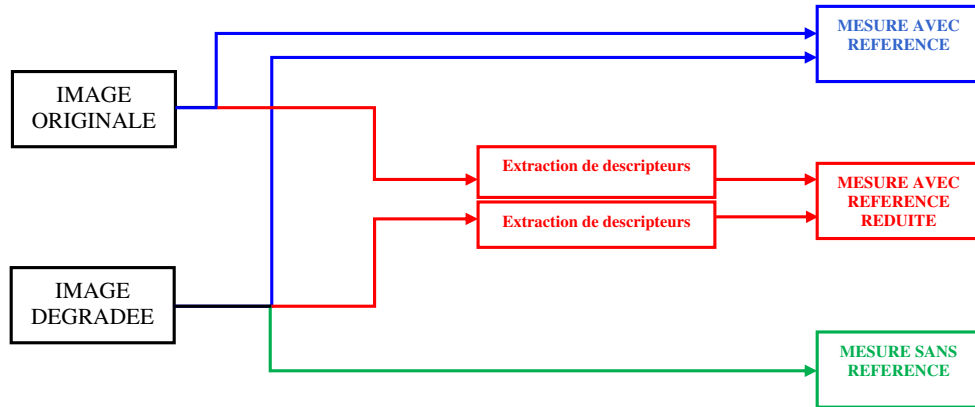


FIGURE 1.14 – Qualité d'image objective : les différentes approches.

structures locales de l'image, les approches neuronales et les mesures inspirées du SVH.

La section suivante est entièrement dédiée à l'évaluation objective. Nous y décrivons les différentes approches proposées ainsi qu'un ensemble représentatif de métriques.

1.4.1 Approches basées EQM

L'erreur quadratique moyenne est l'une des métriques les plus populaires et les plus utilisées à ce jour. Cela est dû essentiellement à sa simplicité, au temps de calcul et surtout au fait que l'on puisse l'intégrer facilement dans un processus d'optimisation tel que la minimisation d'une fonction de coût ou de distorsion mathématique. Cette mesure n'est rien d'autre que la moyenne quadratique du signal erreur ou distorsion. Elle est donnée par :

$$EQM = \frac{1}{M \cdot N} \sum_{i=1}^M \sum_{j=1}^N (I(i, j) - I_d(i, j))^2 \quad (1.5)$$

avec $I(i, j)$ et $I_d(i, j)$ le pixel de coordonnées (i, j) de l'image originale et de sa version dégradée, respectivement. N et M représentent respectivement le nombre de lignes et de colonnes de l'image.

A partir de cette définition, différentes variantes ont été proposées. Le rapport signal à bruit (SNR : Signal Noise Ratio) et le rapport signal à bruit crête (PSNR : Peak Signal Noise Ratio) restent les plus utilisées.

$$SNR = 10\log\left(\frac{\sum_{i=1}^M \sum_{j=1}^N (I(i, j))^2}{EQM}\right) \quad (1.6)$$

$$PSNR = 10\log\left(\frac{\text{Valeur Crête du signal original}}{EQM}\right) \quad (1.7)$$

où la valeur crête du signal originale est fixée à 255 pour les images codées sur 8 bits par pixels.

Depuis les années 80, on n'arrête pas de signaler l'insuffisance de ce type de métriques dans l'évaluation de la qualité d'image à cause de leur faible corrélation avec l'appréciation subjective. Cela n'empêche pas leur utilisation jusqu'à nos jours et surtout dans les applications liées à la vidéo. Un article très récent [WB09]^[1] soulève cette question et dresse un constat favorable aux approches inspirées du SVH. Dans ce travail nous avons essayé de donner notre point de vue sur la question en proposant une approche qui exploite l'existant de façon coopérative et constructive. Nous avons ainsi évité de rentrer dans la logique de compétition frontale "une mesure face à toutes les autres". Nous avons plutôt opté pour une logique d'exploitation de l'existant et de développement d'idées de fusion et d'enrichissement des connaissances dans ce domaine très évolutif. Nous pensons que le nombre de métriques proposées est très suffisant et qu'il est plus judicieux de bien les exploiter plutôt que d'essayer de proposer encore une $n^{ième}$ métrique.

Plusieurs auteurs ont tenté de reprendre les mesures de type EQM et d'améliorer les performances. L'un des premiers modèles a été proposé par Mannos et al. [MS74]^[2]. Ces auteurs proposent d'intégrer au calcul de l'EQM une fonction linéaire et un filtrage par une fonction de sensibilité au contraste (CSF, définie

[1] Z. Wang and A.C. Bovik. Mean squared error love it or leave it - A new look at signal fidelity measures. *IEEE Signal Processing Magazine*, Vol. 26, pp. 98-117, 2009.

[2] J.L. Mannos and D.J. Sakrison. The effects of a visual fidelity criterion on the encoding of images. *IEEE Transactions Information Theory*, Vol. 4, pp. 525-536, 1974.

ultérieurement). Une approche similaire a aussi été présentée dans [MV93]^[1].

Beghdadi et al. proposent quant à eux de reprendre la formule du SNR et de l'appliquer dans l'espace temps-fréquence de Wigner-Ville [BI01]^[2]. Un modèle de masquage simple a aussi été introduit. Le même principe a été repris et étendu au domaine des ondelettes [BPP03]^[3]. Gayle et al. proposent de mesurer la qualité à partir de l'écart-type de l'image originale et de sa version dégradée dans le domaine des ondelettes [GMUE02]^[4].

Une autre mesure intéressante a été proposée par Egiazarian et al. [EAP⁺06]^[5]. Les auteurs proposent d'intégrer la CSF dans le domaine TCD (Transformée en Cosinus Discrète). Une version plus récente incluant aussi un modèle de masquage a été proposée dans [PSE⁺07]^[6].

Chandler et al. proposent une métrique de qualité basée sur l'analyse des coefficients d'ondelettes [CH07b]^[7]. Un seuil de détection perceptuelle de la distorsion est d'abord calculé et comparé à la dégradation. Selon les résultats de comparaison, l'image dégradée est considérée comme visuellement identique ou non à l'originale. Un index de qualité inspiré du SNR est alors calculé.

-
- [1] T. Mitsa and K. Varkur. Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms. *International Conference on Acoustics, Speech, and Signal Processing*, pp. 301-304, 1993.
 - [2] A. Beghdadi and R. Iordache. A wigner-ville distribution-based image dissimilarity measure. *ISSPAInternational Symposium on Signal Processing and Its Applications*, Vol. 1, pp. 430-433, 2001.
 - [3] A. Beghdadi and B. Pesquet-Popescu. A New Image Distortion Measure Based Wavelet Decomposition. *ISSPAInternational Symposium on Signal Processing and Its Applications*, Vol. 2, pp. 485-488, 2003.
 - [4] D. Gayle, H. Mahlab, Y. Ucar and A.M. Eskicioglu. A full-reference color image quality measure in the DWT domain. *European Signal Processing Conference*, 2002.
 - [5] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti and M. Carli. New full-reference quality metrics based on HVS. *Workshop on Video Processing and Quality Metrics*, 2006.
 - [6] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola and V. Lukin. On between-coefficient contrast masking of DCT basis functions. *Workshop on Video Processing and Quality Metrics*, 2007.
 - [7] D. Chandler and S. Hemami. Subjective image database. <http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html>, 2007.

D'autres modèles basés sur l'EQM ont aussi été proposés [Wat93]^[8] [Nil85]^[9] [Hal81]^[10]. Bien que, comparativement aux mesures classiques telles que le PSNR, ces modèles donnent de meilleurs résultats, certaines propriétés importantes du SVH ne sont pas prises en compte. Cela les rend peu corrélées à l'appréciation subjective. Les résultats de corrélation rapportés dans certains travaux sont très bons mais uniquement sur des bases d'images limitées. Nous avons constaté que les corrélations changent d'une base à une autre. Nous sommes aujourd'hui loin de disposer d'une base d'images et de tests acceptés de tous. Certaines bases sont plus adaptées à un type de dégradation et de métriques que d'autres.

1.4.2 Approches structurelles

Pour pallier la faiblesse des mesures type EQM, basées sur le pixel en général, certains auteurs ont orienté leur recherche vers les métriques de qualité basées sur une analyse des structures locales du signal image.

Parmi les métriques les plus populaires et les plus utilisées, on trouve celles proposées par Wang et al., appelées UQI [WB02a]^[11] et SSIM [WBE05]^[12]. Ces mesures sont basées sur l'extraction d'attributs structurels locaux de l'image à partir desquels chaque bloc de l'image est décrit par sa luminance (l), son contraste (c) et sa structure (s). La mesure SSIM est une version améliorée de UQI, elle est donnée par :

-
- [8] A. B. Watson DCtune: A technique for visual optimization of DCT quantization matrices for individual images. *Society for Information Display Digest of Technical Papers*, Vol. 14, pp. 946-949, 1993.
 - [9] N. B. Nill. A visual weighted cosine transform for image compression and quality assessment. *IEEE Transactions on communications*, Vol. 33, pp. 551-556, 1985.
 - [10] C. H. Hall. Subjective evaluation of a perceptual quality metric. *SPIE - Image quality*, Vol. 310, pp. 2000-2004, 1981.
 - [11] Z. Wang and A. Bovik A universal image quality index. *IEEE Signal Processing Letters*, Vol. 2, pp. 81-84, 2002.
 - [12] Z. Wang, A.C. Bovik and E.P. Simoncelli. Structural approaches to image quality assessment. *Handbook of image and video processing*, 2nd Edition, Al Bovik, ed, Academic Press, 2005.

$$SSIM = \frac{1}{W} \sum_{i=1}^W SSIM(i) = \frac{1}{W} \sum_{i=1}^W l(i).c(i).s(i) \quad (1.8)$$

où

$$l(i) = \frac{2.\overline{x(i)}.\overline{y(i)} + C_1}{\overline{x(i)}^2 + \overline{y(i)}^2 + C_1} \quad (1.9)$$

$$c(i) = \frac{2.\sigma_x(i).\sigma_y(i) + C_2}{\sigma_x^2(i) + \sigma_y^2(i) + C_2} \quad (1.10)$$

$$s(i) = \frac{2.\sigma_{xy}(i) + C_3}{\sigma_x(i).\sigma_y(i) + C_3} \quad (1.11)$$

$$C_1 = (L.K_1)^2, C_2 = (L.K_2)^2, C_3 = \frac{C_2}{2} \quad (1.12)$$

où L est la dynamique des niveaux de gris. K_1 et K_2 sont deux constantes fixées par les auteurs à 0.01 et 0.03, respectivement.

Cette mesure a été reprise et améliorée par différents auteurs. Chen et al. proposent de rendre la métrique plus robuste aux dégradations de type flou [CYPX06]^[1], en appliquant le même principe aux points contours de l'image. D'autres variantes intéressantes ont aussi été proposées [CYA06]^[2] [WSB03]^[3] [BP06]^[4] [WS05]^[5]. Au

-
- [1] G. Chen, C. Yang, L. Po and S. XIE. Edge-based structural similarity for image quality assessment. *International Conference in Acoustics, Speech and Signal Processing, Vol. 2*, pp. 933-936, 2006.
 - [2] G. Chen, C. Yang, L. Po and S.L Andwie. Gradient-based structural similarity for image quality assessment. *International Conference on Image Processing, Vol. 2*, pp. 2929-2932, 2006.
 - [3] Z. Wang, E.P. Simoncelli and A.C Bovik Multi-scale structural similarity for image quality assessment. *Asilomar Conference on Signals, Systems and Computers*, 2003.
 - [4] A.C. Brooks and T.N. Pappaas. Structural similarity quality metrics in a coding context: exploring the space of realistic distortions. *SPIE, Human Vision and Electronic Imaging XI, Vol. 6057*, pp. 299-310, 2006.
 - [5] Z. Wang and E. Simoncelli. Translation insensitive image similarity in complex wavelet

vu de son succès, SSIM paraît être une mesure qui tend à remplacer celles basées sur l'EQM.

Shnayderman et al. ont proposé d'analyser localement le signal image à partir de ses valeurs singulières [SGE04]^[6]. Un index de qualité est calculé à partir des éléments de la décomposition en valeurs singulières (SVD) de chaque bloc de l'image.

Ces méthodes ont permis de se rapprocher du comportement du SVH. Ainsi, l'intégration de l'activité locale dans le processus d'estimation de la qualité améliore grandement les performances des systèmes de mesures actuels. Cependant, d'autres aspects importants du SVH doivent être considérés et intégrés de façon à se rapprocher le plus possible de l'appréciation subjective. Ainsi, plusieurs travaux ont été orientés dans ce sens.

1.4.3 Approches inspirées du Système Visuel Humain

Dans cette section, nous décrivons quelques métriques inspirées complètement du SVH. Avant de les présenter, il paraît important de rappeler les caractéristiques principales du SVH.

1.4.3.1 Le Système Visuel humain

Le SVH est un système complexe et pas encore totalement maîtrisé. Néanmoins, il peut être considéré comme un système de transcription d'informations en données exploitable par le cerveau (voir figure 1.15). Une étape de conversion permet de capter l'information reçue et de la transformer en signaux décodables par le cerveau. C'est le rôle de l'oeil qui convertit l'énergie lumineuse en signaux nerveux pour ensuite les transmettre au cortex visuel à travers les nerfs optiques et les corps genouillés. Le cortex visuel décrypte et traite l'information reçue.

domain. *International Conference on Acoustics, Speech and Signal Processing*, Vol. 2, pp. 573-576, 2005.

- [6] A. Shnayderman, A. Gusev and A.M. Eskicioglu. A multidimensional image quality measure using Singular Value Decomposition. *SPIE Image Quality and System Performance Conference*, Vol. 5294, pp. 82-92, 2004.

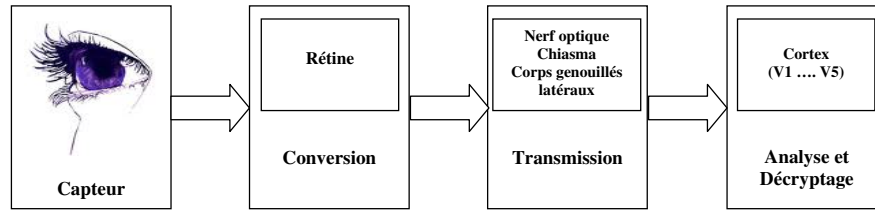


FIGURE 1.15 – Système Visuel Humain : Schéma simplifié.

Afin de mieux comprendre le cheminement de l'information et le fonctionnement du SVH, nous rappelons les éléments essentiels qui entrent en jeu dans le processus de recueil et de traitement de l'information visuelle (capture, conversion, transmission et traitement de l'information).

– Capteur :

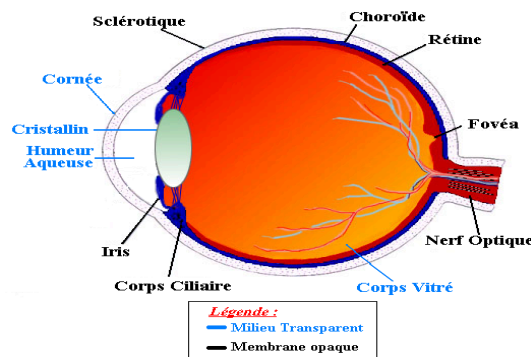


FIGURE 1.16 – Description de l'œil.

L'œil n'est rien d'autre qu'un système optique dont le rôle principal est de faire converger les signaux lumineux vers la zone de conversion et de les transmettre sous forme décodable par le cerveau. Il est constitué de plusieurs éléments importants dont :

La cornée : Du point de vue optique, la cornée est un dioptré sphérique convexe qui sépare l'œil du milieu extérieur. Il a pour rôle principal de concentrer les rayons lumineux reçus vers la rétine.

L'iris : Permet d'adapter l'intensité lumineuse en faisant varier son ouverture. Il joue ainsi le rôle de diaphragme optique.

Le cristallin : Permet de rediriger le flux lumineux vers la rétine où sont disposées les cellules photoréceptrices (les cônes et les bâtonnets). Il joue le rôle de lentille à focale variable.

– **Conversion :**

La rétine joue le rôle de convertisseur d'information. Elle convertit les signaux lumineux capturés par les cellules photoréceptrices en signaux électriques qui sont ensuite transmis au cortex visuel via le nerf optique. La rétine est constituée essentiellement de deux types de récepteurs.

Les cônes : Concentrés autour de la fovéa, leur densité décroît au fur et à mesure que l'on s'éloigne de cette zone. Ils sont au nombre de 6 millions environ et sont sensibles aux détails.

Les bâtonnets : Beaucoup plus nombreux et localisés en majorité en dehors de la zone fovéale, ces récepteurs sont sensibles aux basses luminances (vision floue et grossière) et interviennent plutôt en vision nocturne (monochrome).

– **Transmission :**

Le nerf optique se charge de transporter l'information de la rétine au cortex visuel, en passant par le chiasma et les corps genouillés latéraux.

Le chiasma : Lieu de croisement de l'information provenant de l'oeil droit et de l'oeil gauche, le chiasma a pour rôle de transmettre les informations reçues aux corps genouillés latéraux. Il est à noter que les informations se croisent sans se mélanger.

Les corps genouillés latéraux (CGL) : Constitués de plusieurs catégories de neurones dont les rôles sont bien définis (M, P et K), les corps genouillés (ou géniculés) jouent le rôle de relais de l'information entre le chiasma et le cortex visuel primaire (aire V1).

– **Décryptage :**

Le cortex visuel a pour rôle majeur de décrypter (ou décoder) et d'analyser les signaux reçus. Il est composé de plusieurs parties (aires). L'aire V1 reçoit les signaux des CGL. Une sélection fréquentielle et directionnelle est réalisée à ce niveau. D'autres zones (ou aires) ont été aussi mises en évidence : l'aire V2 (codage couleur, formes, direction du mouvement), l'aire V3 (formes), l'aire V4 (codage des orientations et couleur), l'aire V5 (mouvement directionnel). L'aire V1 reste à ce jour la zone la mieux connue.

1.4.3.2 Les modèles

De nombreuses expériences ont permis de mettre en évidence certaines spécificités du SVH telles que la sensibilité au contraste ou la sélectivité fréquentielle. Suite à ces expériences et constatations, des modèles ont été proposés. Dans cette section, nous présentons ces phénomènes ainsi que les modèles proposés.

– Le contraste :

Le contraste constitue un paramètre important dans les expériences psycho-visuelles. Il est aussi l'une des grandeurs les plus étudiées et les plus utilisées dans les métriques de qualité d'image. Au sens le plus général, il permet de mesurer la sensation relative d'un stimulus par rapport à un autre ou par rapport à un fond dans lequel il est observé. Dans le cas des images de luminance, il peut être quantifié par le rapport entre la luminance locale et la luminance moyenne. Cependant, malgré le nombre considérable d'études sur ce phénomène, il n'existe pas de définition universelle du contraste de luminance. Dans la suite nous rappellerons quelques définitions.

Un des contrastes les plus anciens et les plus connus est sans doute celui de Michelson [Mic27]^[1]. Dans son expérience, Michelson introduit un rapport permettant de quantifier le contraste. Si L_{min} et L_{max} désignent respectivement les valeurs minimum et maximum de la luminance, ce contraste est donné par l'équation 1.13 :

$$C_M = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \quad (1.13)$$

[1] A. Michelson. *Studies in Optic*. Chicago Press, 1927.

L'expérience de Weber-Fechner a introduit une autre définition du contraste [Cor70] [Fra90]. Dans cette expérience, on considère un fond de luminance uniforme L contenant un objet présentant un incrément de luminance ΔL . On s'intéresse alors au seuil de visibilité de l'objet quand on fait varier la luminance du fond. Il s'agit d'étudier la relation entre l'incrément ΔL , nécessaire pour rendre juste visible l'objet, et la luminance L du fond uniforme. Pour étudier cette dépendance on trace le rapport $\Delta L/L$ en fonction de L . Le rapport ainsi défini est appelé contraste de Weber-Fechner, et le seuil de visibilité plus communément connu sous le nom de JNC (Just-Noticeable Contrast) est donné par l'équation 1.14 :

$$C_W = \frac{\Delta L}{L} \quad (1.14)$$

La figure 1.17 représente l'image test utilisée dans l'expérience de Weber-Fechner ainsi que le tracé log-log de la variation de l'incrément de luminance juste perceptible ΔL en fonction de la luminance du fond L . On observe une première loi de variation, dans une zone de faible étendue appelée zone de De Vries-Rose, de pente 0.5 suivie de la zone bien connue de Weber-Fechner (W-F) de pente 1, et enfin une zone de saturation de pente 2. En général, seule la zone W-F est exploitée en traitement d'image en dehors de quelques travaux intéressants où les autres zones sont considérées [KP86]^[2] [TJZ96]^[3].

Le contraste de Weber-Fechner est resté très longtemps la seule référence dans le domaine jusqu'au début des années 40. Moon et Spencer, en s'appuyant sur les travaux de Holladay et ceux de Hecht [MS43]^[4] [MS44]^[5], ont pu étendre la notion de contraste de luminance au cas d'objet noyé dans un fond de luminance non uniforme. L'idée principale de Moon et Spencer est

-
- [2] M. K. Kundu and S. K. Pal. Thresholding for edge detection using human psychovisual phenomena. *Pattern Recognition Letters*, Vol. 4, pp. 433-441, 1986.
 - [3] Z. Tianxu, P. Jiaxiong and L. Zongjie. An adaptive image segmentation method with visual nonlinearity characteristics. *IEEE Transactions on Systems, Man Cybern.*, Vol. 26, pp. 619-627, 1996.
 - [4] P. Moon and D. E. Spencer. The specification of foveal adaptation. *Journal of the Optical Society of America*, Vol. 33, pp. 444-456, 1943.
 - [5] P. Moon and D. E. Spencer. Visual data applied to lighting design. *Journal of the Optical Society of America*, Vol. 34, pp. 230-240, 1944.

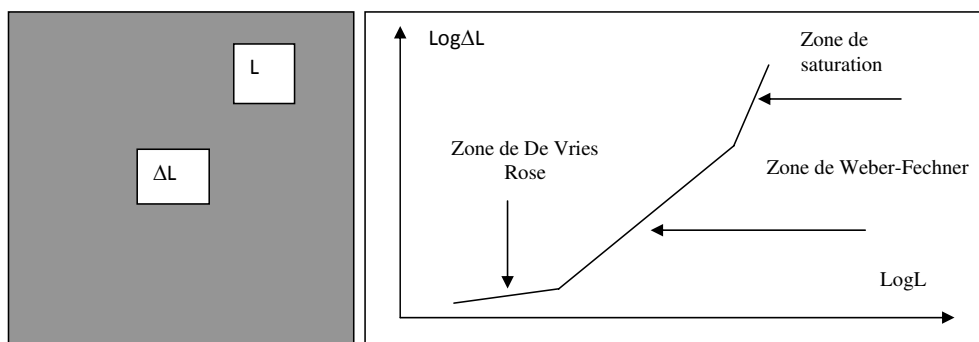


FIGURE 1.17 – Expérience de Weber-Fechner et tracé de la loi de variation du seuil de perception.

d'appliquer le principe de Holladay selon lequel tout fond non uniforme peut être remplacé par un autre de luminance uniforme et produisant le même effet au niveau de la perception. Cela conduit à la définition de la luminance d'adaptation.

Le calcul de la luminance d'adaptation L_A , effectué par Moon et Spencer, conduit à une expression où interviennent la luminance L_C du voisinage immédiat ou couronne et celle du fond lointain notée L_F . La relation liant ces trois quantités est :

$$L_A = \alpha_C L_C + \alpha_F L_F \quad (1.15)$$

où les coefficients α_C et α_F ($\alpha_C = 0.923$, $\alpha_F = 0.077$), correspondent aux contributions du voisinage immédiat et du fond lointain, respectivement. La représentation géométrique simplifiée du modèle de l'image fovéale de Moon-Spencer est donnée par la figure 1.18.

Une autre mesure de contraste inspirée de celle de Gordon et Rangayan [GR84]^[1] et qui a connu un réel succès, notamment pour le rehaussement de contraste, est celle de Beghdadi et al. [BN89]^[2]. Cette définition de contraste a

-
- [1] R. Gordon and R. M. Rangayan. Feature enhancement of film mammograms using fixed and adaptive neighborhood. *Journal of the Optical Society of America*, Vol. 23, pp. 560-564, 1984.
 - [2] A. Beghdadi and A. Le Négrate Contrast enhancement technique based on local detection

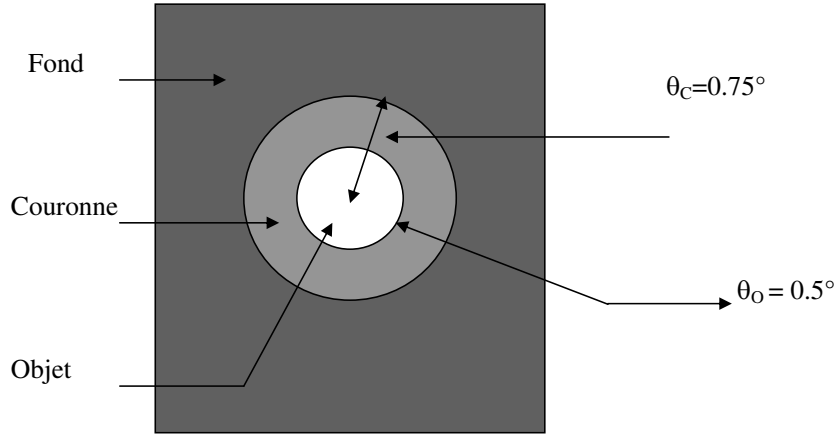


FIGURE 1.18 – Représentation 2D de l'image fovéale dans le modèle de Moon-Spencer.

été reprise telle quelle ou améliorée dans plusieurs travaux [KA90]^[3] [KPSP97]^[4]. Elle s'appuie sur le fait que le SVH est plus sensible aux contours qu'aux autres structures de l'image. Ainsi, ce contraste est défini à partir du niveau moyen des contours estimé dans une fenêtre d'analyse mobile. En chaque pixel de coordonnées (k,l) , centre de la fenêtre d'analyse W_{kl} de taille impaire, on estime le niveau de luminance des contours en utilisant l'estimateur donné par :

$$\overline{E_{kl}} = \frac{\sum_{(i,j) \in W_{kl}} \varphi(\Delta_{ij}) f_{ij}}{\sum_{(i,j) \in W_{kl}} \varphi(\Delta_{ij})} \quad (1.16)$$

Où f_{ij} est le niveau de gris au point (i,j) et $\varphi(\Delta_{ij})$ est une fonction croissante du module du gradient Δ_{ij} en (i,j) .

of edges. *Computer Vision, Graphics, and Image Processing*, pp. 162-174, 1989.

- [3] F. Kammoun and J.P. Astruc. Augmentation de contraste adaptative suivant les niveaux de gris des contours. *Société française de physique, Vol. 25*, pp. 1029-1035, 1990.
- [4] J. K. Kim, J. M. Park, K. S. Song and H. W. Park. Adaptive mammographic image enhancement using first derivative and local statistics. *IEEE Transactions on Medical Imaging, Vol. 16*, pp. 495-502, 1997.

Un exemple de fonction simple est l'élévation à une puissance entière du module du gradient. La robustesse de cet estimateur a été étudiée plus en détail par Wilkinson [Wil98]^[1]. Il a été aussi démontré que le pseudo-gradient obtenu par combinaison de deux réponses de l'opérateur de Sobel, pour calculer Δ_{ij} , fournit l'estimateur le plus robuste vis à vis du bruit blanc gaussien additif. Une fois ce niveau estimé, on calcule le contraste local associé au pixel (k,l) à partir de l'expression :

$$C_{kl} = \frac{|\overline{E_{kl}} - f_{k,l}|}{\overline{E_{kl}} + f_{kl}} \quad (1.17)$$

Partant de cette définition et en s'inspirant des travaux de Burt et Adelson [BA83]^[2], Jolion a introduit la notion de contraste multirésolution ou contraste pyramidal [Jol94]^[3]. L'image est analysée au moyen de la décomposition multirésolution de type pyramide Gaussienne. A chaque niveau de résolution (j) , le contraste local au point de coordonnées (k,l) , centre de la fenêtre d'analyse W_{kl} , est défini par :

$$C_{kl}^{(j)} = \frac{g_{kl}^{(j)} + e_{kl}^{(j)}}{g_{kl}^{(j)} - e_{kl}^{(j)}} \quad (1.18)$$

Où $g_{kl}^{(j)}$ est le niveau de gris à la résolution (j) et $e_{kl}^{(j)}$ est le signal issu de la pyramide Laplacienne.

L'intérêt d'une telle mesure est de pouvoir analyser le contraste à différentes échelles de résolution simulant ainsi l'effet d'éloignement de l'observateur de l'image observée. On peut aussi citer une autre définition du contraste similaire qui est basée sur la transformation en ondelettes. Cette mesure permet

-
- [1] M.H.F. Wilkinson. Optimizing Edge Detectors for Robust Automatic Threshold Selection: Coping with Edge Curvature and Noise. *Graphical models and image processing*, Vol. 60, pp. 385-401, 1998.
 - [2] P. J. Burt and H. E. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communication*, Vol. 31, pp. 532-540, 1983.
 - [3] J-M. Jolion. Analyse multirésolution du contraste dans les images numériques. *Traitement du Signal*, Vol. 11, pp. 245-255, 1994.

d'introduire la notion de contraste multi-échelle et fournit le moyen d'analyser l'image à différents niveaux de détails [AMPD91]^[4].

En se basant sur la sensibilité fréquentielle du SVH, Peli définit un contraste par bande de fréquences spatiales [Pel90]^[5]. L'image est décomposée en plusieurs canaux au moyen d'un banc de filtres passe bande de type Gabor, par exemple, selon l'opération de convolution :

$$g_k(m, n) = f(m, n) * h_k(m, n) \quad (1.19)$$

Où h_k est la réponse impulsionnelle du filtre correspondant au $k^{\text{ième}}$ canal et g_k est la version filtrée du signal original f . Le contraste au point (m, n) , et dans la bande k , est alors égal au rapport de la luminance dans la bande k et de la composante g_O du fond continu en ce même point.

$$C_k(m, n) = \frac{g_k(m, n)}{g_O(m, n)} \quad (1.20)$$

Ce contraste est bien adapté aux images complexes et est souvent utilisé dans le calcul de certaines métriques de qualité d'image [OMM97]^[6]. Notons que Peli ne fait pas intervenir l'aspect directionnel et utilise de ce fait un filtre passe-bande isotrope.

Une autre définition du contraste conduisant à deux expressions a été pro-

-
- [4] J-P. Antoine, R. Murenzi, B. Piette and M. Duval. Image analysis with 2D continuous Wavelet Transform Detection of position, orientation and visual contrast of simple objects. *Wavelets and Applications*, Y. Meyer Editor, Masson/Springer-Verlag, Vol. 11, pp. 144-159, 1991.
 - [5] E. Peli. Contrast in complex images. *Journal of the Optical Society of America*, Vol. 7, No. 10.
 - [6] W. Osberger, A.J. Maeder and D. McLean. A Computational Model of the Human Visual System for Image Quality Assessment. *Digital Imaging Computing : Techniques Applications*, pp. 337-342, 1997.

posée par Lillesaeter [Lil93]^[1]. La première ne fait intervenir que la luminance et la seconde, difficilement exploitable en pratique, intègre la forme de l'objet perçu. Partant de l'expression de Weber-Fechner, et ayant remarqué l'asymétrie de cette mesure ainsi que le fait que les contrastes de Weber-Fechner négatifs et positifs de même valeur absolue ne sont pas perçus de la même manière par le SVH, Lillesaeter propose l'expression suivante :

$$C = \log\left(\frac{L_O}{L_F}\right) \quad (1.21)$$

Où L_O et L_F désignent respectivement la luminance moyenne de l'objet et celle du fond.

Notons que cette mesure de contraste est équivalente à celle de Weber-Fechner dans le cas où L_O et L_F sont très voisins (à 2% près). $C = \frac{L_O - L_F}{L_F} \approx \log(L_O) - \log(L_F)$ (si $|C| \ll 1$). Une deuxième expression où la géométrie des contours de l'objet perçu est intégrée a aussi été proposée dans [Lil93]^[1]. Cette dernière mesure n'est toutefois pas facilement utilisable en pratique car elle nécessite au préalable une détection des contours de l'objet et un calcul d'intégrale le long des frontières ainsi déterminées.

Plusieurs définitions du contraste dans les images couleur ont aussi été proposées [ARM04]. Parmi elles citons, celle proposée par Favier et al. [FT99]^[2] qui est basée sur une analyse par régions adjacentes. Les auteurs proposent de combiner une mesure d'émergence et de contraste couleurs de contours, obtenus à partir d'une image segmentée. La mesure d'émergence est dérivée d'un graphe d'adjacence. La mesure de contraste couleur de contour est quant à elle obtenue à partir d'une analyse des contours inter-régions. Cette approche permet de prendre en compte certaines propriétés du SVH liées notamment à la taille (en termes de surface) des régions adjacentes et à la forme des régions.

Nous tenons à rappeler qu'il n'existe pas de définition universelle et qu'à une

-
- [1] O. Lillesaeter. Complex contrast, a definition for structured targets and backgrounds. *Journal of the Optical Society of America*, Vol. 10, pp. 2453-2457, 1993.
 - [2] E. Favier and A. Trémeau. Une mesure de contraste couleur inter-régions. *GRETSI*, pp. 643-646, 1999.

application donnée, il convient de choisir celle qui est la plus adaptée.

– **La sensibilité aux contrastes :**

Le modèle basé sur la sensibilité au contraste à partir de stimuli simples permet de définir des fonctions de sensibilité au contraste (CSF : Contrast Sensitivity Function) en fonction de la fréquence spatiale ou temporelle. Il est à noter que dans la plupart des expériences, le contraste de Michelson a été utilisé.

Pour un stimulus sinusoïdal de luminance L , le contraste peut être défini comme le rapport de l'amplitude de la variation sinusoïdale sur la luminance moyenne. Le contraste minimum requis pour la détection du stimulus est appelé seuil de contraste (ou contraste au seuil de visibilité). Comme la mesure de la sensibilité au contraste s'effectue en fonction de variations sinusoïdales de la luminance, la sensibilité au contraste de l'oeil humain est généralement définie comme l'inverse du seuil de contraste. De plus, le seuil de contraste dépend plus généralement de la longueur d'onde de la variation sinusoïdale de la luminance, c'est-à-dire de la distance entre les maximums des sinusoïdes.

En général, la fonction de sensibilité au contraste est plus souvent exprimée comme une fonction de la fréquence spatiale. Notre oeil est beaucoup plus sensible aux basses fréquences et assez limité aux hautes fréquences, comme on peut le constater dans la figure 1.19 où chaque courbe montre la sensibilité au contraste avec des luminances moyennes différentes de $0,01\text{cd}/\text{m}^2$ à $1000\text{cd}/\text{m}^2$. La CSF peut aussi dépendre de la fréquence temporelle (variation temporelle de la luminance).

Différents modèles de fonction CSF sont cités dans [PAYG93]^[3] [WYSV97]^[4]. Dans la suite, nous présentons quelques fonctions CSF proposées dans la littérature.

-
- [3] E. Peli, L.E. Arend, G.M. Young and R.B. Goldstein. Contrast Sensitivity to Patch Stimuli: effects of Spatial Bandwidth and Temporal Presentation. *Spatial Vision*, Vol. 7, pp. 1-14, 1993.
 - [4] A.B. Watson, G.Y. Yang, J.A. Solomon and J Villaseñor. Visibility of wavelet quantisation noise. *IEEE Transactions on Image Processing*, Vol. 6, pp. 1164-1175, 1997.

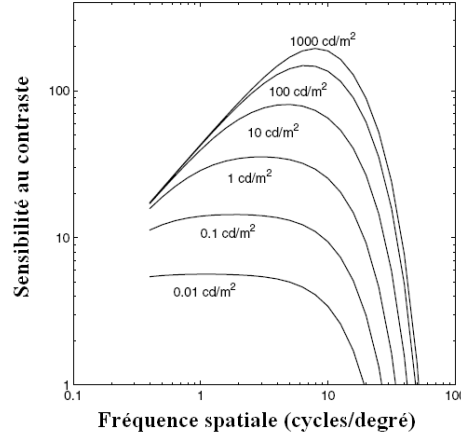


FIGURE 1.19 – Fonction de sensibilité au contraste (CSF).

CSF de Mannos et Sakrison

En prenant en compte non seulement la luminance mais aussi l'orientation dans le domaine fréquentiel, Mannos et Sakrison proposent la fonction CSF suivante [MS74]^[1] :

$$CSF(f) = a(b + \frac{f}{f_O}) \exp(-\frac{f}{f_O})^c \quad (1.22)$$

où a , b , c et f_O sont des paramètres du modèle et $f = \sqrt{u^2 + v^2}$ où (u, v) est la direction dans le domaine fréquentiel en cycles par degré.

En particulier, l'expression $CSF(f) = 2,6.(0,192 + 0,114.f).exp^{(-0,114.f^{1,1})}$ a été utilisée par les auteurs, dans le cadre de l'évaluation de la qualité des images dégradées.

CSF de Barten

Barten propose une fonction CSF [Bar90]^[2] basée sur la variation de luminance du fond L et des conditions d'observation. Elle est donnée par l'ex-

-
- [1] J.L. Mannos and D.J. Sakrison. The effects of a visual fidelity criterion on the encoding of images. *IEEE Transactions Information Theory*, Vol. 4, pp. 525-536, 1974.
 - [2] P.G.J. Barten. Evaluation of subjective image quality with the square-root integral method. *Vision Research*, Vol. 7, pp. 2024-2031, 1990.

pression suivante :

$$CSF(f, L, \omega) = a(L, f, \omega) \cdot f \cdot \sqrt{1 + 0,006 \cdot \exp^{b(L) \cdot f}} \cdot \exp^{-b(L) \cdot f} \quad (1.23)$$

avec

$$a(L, f, \omega) = \frac{540 \cdot (1 + 0,7/L)^{-0,2}}{1 + \frac{12}{\omega(1+f/3)^2}} \quad (1.24)$$

$$b(L) = 0,3 \cdot (1 + 100/L)^{0,15} \quad (1.25)$$

L'angle ω (exprimé en degré) est donné par $\omega = \frac{180 \cdot \sqrt{A}}{\pi \cdot D}$ où A est la taille de l'image et D la distance d'observation. Cette CSF a été modifiée pour rendre compte du caractère anisotrope de la sensibilité visuelle :

$$CSF(f, L, \omega, \theta) = a(L, f, \omega) \cdot f \cdot \sqrt{1 + 0,006 \cdot \exp^{b(L) \cdot f}} \cdot \exp^{-b(L) \cdot f \cdot \Gamma(\theta)} \quad (1.26)$$

avec

$$\Gamma(\theta) = 1 - 0,079 \cdot [\cos(4 \cdot \theta) - 1] \quad (1.27)$$

CSF de Dooley

Kundur et al. [KH97]^[3] ont présenté le modèle de Dooley avec une fonction CSF à deux dimensions :

$$C(u, v) = 5,05 \cdot \exp -0,178 \cdot (u + v) (\exp 0,1 \cdot (u + v) - 1) \quad (1.28)$$

où u et v représentent les fréquences.

CSF de Daly

-
- [3] D. Kundur and D. Hatzinakos. A Robust Digital Image Watermarking Method using Wavelet-Based Fusion. *International Conference Image Processing*, pp. 544-547, 1997.

Daly propose une fonction de sensibilité au contraste CSF_D qui dépend de divers paramètres (fréquence radiale, orientation, luminance, surface, distance d'observation et excentricité). Elle est donnée par la relation suivante :

$$CSF_D(\omega, \theta) = P \cdot \min\left\{S\left(\frac{\omega}{r_a \cdot r_e \cdot r_\theta}, L_a, s\right), S(\omega, L_a, s)\right\} \quad (1.29)$$

où P est la sensibilité maximale qui peut varier d'un observateur à un autre (ici $P = 250$). Les paramètres r_a , r_e et r_θ permettent la prise en compte des changements de la largeur de bande en fonction de la distance d'observation, de l'excentricité et de l'orientation, respectivement.

$$r_a = 0,856 \cdot d^{0,14}, \quad r_e = \frac{1}{1 + 0,24 \cdot e_c}, \quad r_\theta = 0,11 \cdot \cos(4 \cdot \theta) + 0,89 \quad (1.30)$$

La fonction S est déterminée par :

$$S(\omega, L_a, s) = ((3,23 \cdot (\omega^2 \cdot s)^{-0,3})^5 + 1)^{1/5} \cdot 0,9 \cdot A_l \cdot \omega \cdot \exp^{-0,9 \cdot B_l \cdot \omega} \cdot \sqrt{1 + 0,06 \cdot \exp^{0,9 \cdot B_l \cdot \omega}} \quad (1.31)$$

où

$$A_l = 0,801 \cdot \left(1 + \frac{0,7}{L_a}\right)^{-0,2}, \quad B_l = 0,3 \cdot \left(1 + \frac{100}{L_a}\right)^{0,15} \quad (1.32)$$

Les modèles de CSF établis à partir de stimuli simples sont valides dans le cas où il n'y a pas d'interaction entre les composantes fréquentielles. Dans la pratique, cette hypothèse n'est pas toujours vraie (comportement non linéaire du SVH). Il est alors important de considérer les effets de masquage qui se manifestent par la variation du seuil de visibilité d'un stimulus en fonction de la présence d'un autre stimulus.

– Sélectivité fréquentielle et directionnelle :

Plusieurs expériences ont mis en évidence le caractère sélectif du SVH (au

niveau du cortex) [Sak77]^[1]. Il est admis que chaque type de cellule du cortex visuel primaire traite uniquement les informations d'une bande de fréquences et orientation données.

Le principe général consiste à décomposer l'image en plusieurs canaux perceptuels imitant ainsi les cellules du cortex. Cette décomposition s'effectue en général par des transformations linéaires. Dans la littérature, il existe plusieurs travaux relatifs à ces modèles multicanaux parmi lesquels on peut citer la transformée Cortex [A.B87]^[2]. Elle est obtenue à partir de la combinaison de filtres à sélectivité radiale (Dom) et angulaire (Fan), comme illustré par la figure 1.20. Cette dernière a été très utilisée et reprise par différents auteurs [Dal93]^[3] [Lub95]^[4].

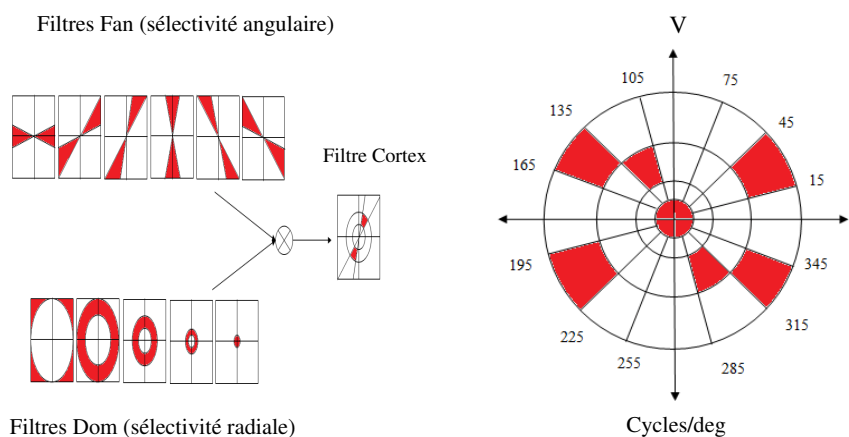


Illustration des découpages en sous-bandes

FIGURE 1.20 – Transformée en Cortex.

-
- [1] D.J. Sakrison. On the role of the Observer and a Distortion Measure in Image Transmission. *IEEE Transactions on Communications*, Vol. 25, pp. 1251-1267, 1977.
 - [2] A.B.Watson. The cortex transform: Rapid computation of simulated neural images. *Computer Vision Graphics and Image Processing*, Vol.39, pp. 311-327, 1987.
 - [3] S. Daly. The visible differences predictor: an algorithm for the assessment of image fidelity. *Digital images and human vision*, MIT Press, 1993.
 - [4] J. Lubin. A visual Discrimination model for imaging system design and evaluation. *Vision model for target detection and recognition*, Eli Peli, Editor, World Scientific, pp. 245-283, 1995.

D'autres modèles similaires ont été proposés, tels que les différences de gaussiennes (DoG) [WG84]^[1], ou les filtres de Gabor [Mar80]^[2]. Dans le cadre de cette thèse, nous nous sommes focalisés essentiellement sur la transformée en Cortex.

– **Le masquage :**

Le phénomène de masquage fait référence aux changements de visibilité d'un stimulus (ou signal masqué) en fonction de la présence d'un autre (masquant) [LF80]^[3]. La visibilité du stimulus peut être soit diminuée (moins visible), on parle alors d'effet de masquage ou au contraire, elle peut être augmentée (plus visible), on parle alors d'effet de facilitation. Ce phénomène est accentué quand le signal masqué et le signal masquant sont d'orientations et de fréquences voisines.

La notion de masquage est donc liée à l'orientation et à la fréquence des signaux que l'on observe. Dans la littérature, on distingue différents types de masquage : le masquage de luminance (adaptation de la luminance) et le masquage de contraste.

Le masquage de luminance :

Cet effet de masquage est lié à la capacité de l'oeil à adapter sa sensibilité à la luminance moyenne globale de l'image. Ce masquage est encore appelé "adaptation de luminance". Par exemple, si le fond est très clair ou très foncé, l'oeil distingue difficilement des objets. Par contre, si celui-ci est de niveau moyen, sa détection est plus facile.

– ***Modèle d'adaptation de luminance de Chou et Li***

Chou et al. ont proposé un modèle permettant de caractériser la variation du seuil de visibilité en fonction de la luminance de fond de l'image

-
- [1] H. R. Wilson and D. J. Gelb. Modified line-element theory for spatial-frequency and width discrimination. *Journal of the Optical society of America*, pp. 124-134, 1984.
 - [2] S. Marcelja. Mathematical description of the response of simple cortical cells. *Journal of the Optical society of America*, Vol. 70, pp. 1297-1300, 1980.
 - [3] G. Legge and J. Foley. Contrast masking in human vision. *Physics Education*, Vol. 70, pp. 1458-4471, 1980.

[CL95]^[4] :

$$f(L_{bg}(x, y)) = \begin{cases} T_O \cdot (1 - \sqrt{\frac{L_{bg}(x, y)}{127}}) + 3 & \text{si } L_{bg}(x, y) \leq 127 \\ \gamma \cdot (L_{bg}(x, y) - 127) + 3 & \text{si } L_{bg}(x, y) > 127 \end{cases} \quad (1.33)$$

où $L_{bg}(x, y)$ est la luminance moyenne du fond dans un voisinage réduit. T_O et γ sont des constantes qui dépendent des conditions de visualisation.

Masquage de contraste :

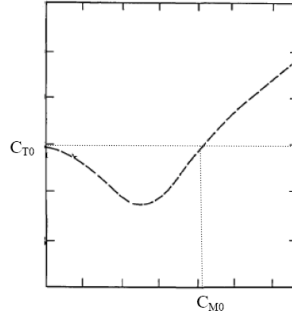


FIGURE 1.21 – Modèle de masquage de contraste.

Pour évaluer le masquage de contraste, on mesure le seuil de détection d'un stimulus lorsqu'il est masqué par un autre stimulus dont le contraste varie. La figure 1.21 représente la courbe de variation du contraste C_T minimum nécessaire pour que le stimulus masqué soit visible (seuil de détection du signal masqué en présence du masque) en fonction du contraste C_M du stimulus masquant. Cette courbe est obtenue pour une fréquence et une orientation fixes. On note alors C_{T0} , le seuil de détection du stimulus (à masquer) en l'absence du masque. La valeur C_{M0} représente la transition entre les deux états (présence et absence de masquage). On s'aperçoit alors :

[4] C.H. Chou and Y.C. Li . A perceptually tuned subband image coder based on the measure of just-noticeable distortion profile. *IEEE Transactions on Circuits and Systems for Video Techniques*, Vol. 5, pp. 467-476, 1995.

- Tant que le contraste C_M du signal masquant est inférieur à C_{TO} , il n'y a pas d'effet de masquage ; on obtient un effet négatif lorsque C_M est proche de C_{TO} .
- Pour un contraste C_M du signal masquant supérieur à C_{TO} , le seuil de détection du stimulus masqué augmente avec le contraste du masque. Autrement dit, le masquage augmente le seuil de visibilité.

On peut aussi constater que le masquage est maximal dans les hautes fréquences. Dans le contexte du tatouage d'image par exemple, la stratégie pour certaines applications (selon la robustesse) est donc d'insérer le tatouage dans les hautes fréquences en utilisant le phénomène de masquage de contraste pour mieux cacher un maximum d'informations tout en maintenant un niveau d'imperceptibilité requis.

– **Modèle de masquage de contraste de Legge et Foley**

Legge et Foley [LF80]^[1] ont proposé un modèle de masquage non linéaire simple, établi pour les structures sinusoïdales à fréquence radiale et orientation fixes (f, θ) . Ce modèle décrit la variation du seuil de sensibilité de contraste (seuil de détection) C_T d'un signal en fonction du contraste C_M du signal masquant (voir figure 1.22).

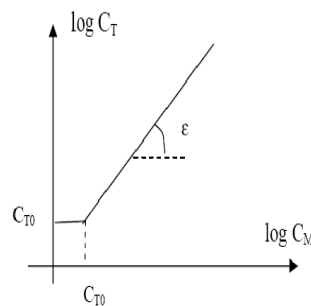


FIGURE 1.22 – Modèle de masquage de contraste.

[1] G. Legge and J. Foley. Contrast masking in human vision. *Physics Education*, Vol. 70, pp. 1458-4471, 1980.

Le modèle établi est alors résumé comme suit :

$$C_{T(f,\theta)}(C_M) = \begin{cases} C_{T_o} & \text{si } C_M \leq C_{T_o} \\ C_{T_o} \cdot (\frac{C_M}{C_{T_o}})^\varepsilon & \text{si } C_M \geq C_{T_o} \end{cases} \quad (1.34)$$

Une extension de ce modèle intégrant les paramètres fréquence et orientation est donnée par :

$$C_T = C_{T_o} + K_{(f_o,\theta_o)} \cdot |C_{T(f_o,\theta_o)}(C_M) - C_{T_o}| \quad (1.35)$$

où $K_{(f_o,\theta_o)}$ est un facteur de couplage en fréquence f_o et orientation θ_o du masque :

$$K_{(f_o,\theta_o)} = \exp \left[- \left(\frac{\log^2(\frac{f}{f_o})}{F^2(f_o)} + \frac{\theta - \theta_o}{\Theta^2(f_o)} \right) \right] \quad (1.36)$$

$F(f_o)$ et $\Theta(f_o)$ sont des paramètres dépendant d'une fonction gaussienne.

On démontre que pour un C_M donné, le masquage $C_T(f, \theta)$ est maximum si les deux signaux ont la même fréquence radiale ($f = f_o$) et la même orientation ($\theta = \theta_o$).

Pour plus de détails, les lecteurs sont invités à se référer à [TH94]^[2] [WS97]^[3].

Cette brève description du SVH et des modèles associés montre combien il est difficile d'intégrer ces quelques connaissances, assez limitées, sur les mécanismes de la vision humaine dans un système de mesure de la qualité d'image. Dans ce qui suit, nous présentons les principales métriques intégrant ces modèles.

-
- [2] P. C. Teo and D. J. Hegger. Perceptual image distortion. *Human Visual Procesing and Digital Display*, Vol. 2179, pp. 127-141, 1994.
 - [3] A.B. Watson and J.A. Solomon. Model of Visual Contrast Gain Control and Pattern Masking. *Journal of th Optical Society of America*, Vol. 14, 2379-2391, 1997.

1.4.3.3 Métriques inspirées du SVH

Parmi les modèles de référence, celui proposé par Daly [Dal93]^[1] appelé VDP (Visual Difference Predictor), est le plus complet et le mieux détaillé. Il intègre toutes les caractéristiques essentielles du SVH qui interviennent dans l'analyse et l'extraction du contenu visuel d'une image. Le principe de ce modèle est d'extraire une carte de visibilité de la dégradation où à chaque point est associé une probabilité de visibilité (voir figure 1.23).

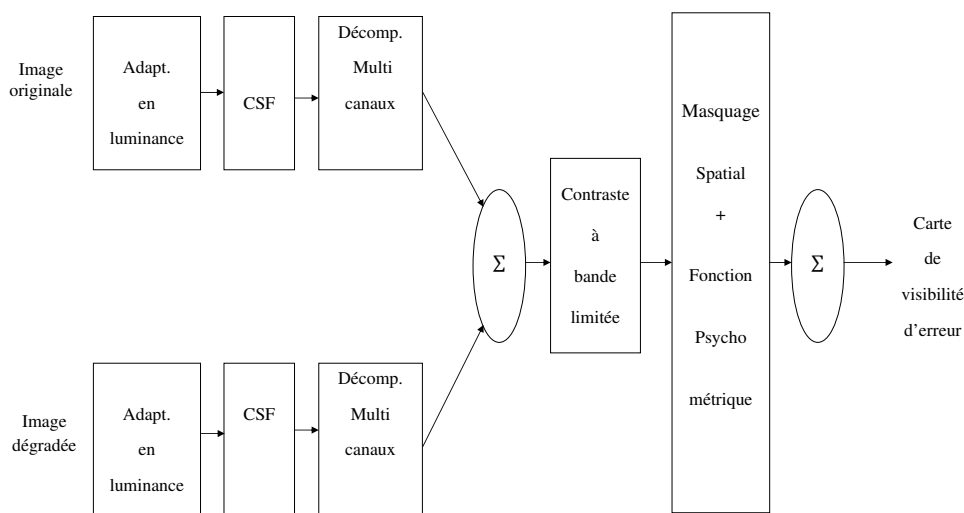


FIGURE 1.23 – VDP : Modèle de Daly.

Une fonction permettant de simuler la sensibilité des cellules de la rétine à la luminance est d'abord appliquée. Daly propose la fonction suivante :

$$L_a(x, y) = \frac{R(x, y)}{R_{max}} = \frac{L(x, y)}{L(x, y) + [12.6 \cdot L(x, y)]^{0.63}} \quad (1.37)$$

où $L(x, y)$ représente l'intensité du pixel (x, y) de l'image.

Les deux images, originale et dégradée, sont ensuite filtrées par une fonction

[1] S. Daly. The visible differences predictor: an algorithm for the assessment of image fidelity. *Digital images and human vision*, MIT Press, 1993.

de sensibilité au contraste. Daly propose d'utiliser la CSF suivante.

$$CSF_D(\omega, \theta) = P.min\{S(\frac{\omega}{r_a.r_e.r_\theta}, L_a, s), S(\omega, L_a, s)\} \quad (1.38)$$

avec

$$S(\omega, L_a, s) = ((3,23.(\omega^2.s)^{-0,3})^5 + 1)^{1/5} \cdot 0,9.A_l.\omega.\exp^{-0,9.B_l.\omega}.\sqrt{1 + 0,06.\exp^{0,9.B_l.\omega}} \quad (1.39)$$

$$r_a = 0,856.d^{0,14}, \quad r_e = \frac{1}{1 + 0,24.e_c}, \quad r_\theta = 0,11.cos(4.\theta) + 0,89 \quad (1.40)$$

$$A_l = 0,801.(1 + \frac{0,7}{L_a})^{-0,2}, \quad B_l = 0,3.(1 + \frac{100}{L_a})^{0,15} \quad (1.41)$$

avec P la valeur maximale de la CSF, L_a la luminance d'adaptation (cd/m^2), s la taille de l'image exprimée en degrés visuels, d la distance d'observation et e_c l'excentricité (en degrés).

La sélectivité fréquentielle et directionnelle du SVH est ensuite modélisée par une décomposition multi-canal. Daly propose d'utiliser la transformée en Cortex avec une décomposition en 5 bandes de fréquences et 6 orientations. Il en résulte donc 31 sous-images (en comptant la bande centrale) obtenues comme suit :

$$m_{k,l}(x, y) = TF^{-1}.\{TF\{L_w(x, y)\}.CSF(u, v).cortex_{k,l}(u, v)\} \quad (1.42)$$

où $m_{k,l}(x, y)$ est la sous image de contraste de la sous bande (k, l) .

Un modèle de masquage est aussi intégré. L'auteur propose de prendre en compte uniquement le masquage intra-canal et introduit la notion de masquage mutuel selon le modèle suivant :

$$T_{k,l}^{em}(x, y) = \min\{T_{k,l}^{e(i)}(x, y), T_{k,l}^{e(\tilde{i})}(x, y)\} \quad (1.43)$$

où

$$T_{k,l}^{e(i)}(x, y) = (1 + (k_1 \cdot (k_2 \cdot |m_{k,l}(x, y)|)^s)^b)^{1/b} \quad (1.44)$$

$T_{k,l}^{e(i)}(x, y)$ et $T_{k,l}^{e(\tilde{i})}(x, y)$ représentent les seuils d'élévation du pixel (x, y) de la sous image contraste $m_{k,l}$ de l'image originale et de sa version dégradée, respectivement. b , k_1 , k_2 et s sont des paramètres fixés par l'auteur. $T_{k,l}^{em}$ est le seuil d'élévation du pixel (x, y) caractérisant le masquage mutuel.

Pour chaque sous-image, une carte d'erreurs perceptuelles est ainsi obtenue. Ces cartes sont ensuite combinées pour décrire la visibilité de la dégradation pour chaque pixel de l'image.

Dans [CSB99]^[1], le modèle de Daly a été repris et amélioré. Dans [MMS04]^[2], Mantiuk et al. ont proposé une extension de VDP pour des images HDR (High Dynamic Range). Une version dans le domaine des ondelettes a aussi été proposée [Bra99]^[3].

Un autre modèle intéressant basé sur le seuil de différence juste visible (JND) a été proposé par Lubin [Lub95]^[4]. Bolin et al. ont proposé une version simplifiée

-
- [1] P. Le Callet, A. Saadane and D. Barba. Orientation selectivity of opponent-colour channels. *European Conference on Visual Perception, Vol. 28, pp. 67*, 1999.
 - [2] R. Mantiuk, K. Myszkowski and H.P. Seidel. Visible difference predictor for high dynamic range images. *International Conference on Systems, Man and Cybernetics, Vol. 3, pp. 2763-2769*, 2004.
 - [3] A. P. Bradley. A wavelet Visible Difference Predictor. *IEEE Transactions on Image Processing, Vol. 8, pp. 717-730*, 1999.
 - [4] J. Lubin. A visual Discrimination model for imaging system design and evaluation. *Vision*

de VDM pour les images de synthèse en utilisant la transformée en ondelettes [BM98]^[5]. Cette méthode a été étendue à la couleur [BM99]^[6]. Une méthode prenant en compte la couleur a aussi été proposée dans [Lub97a]^[7] [Lub97b]^[8].

Dans [HT95]^[9], une mesure de fidélité perceptuelle a été proposée. Le modèle intègre aussi la sensibilité au contraste, le masquage de luminance et de contraste. D'autres modèles intéressants et basés sur les mêmes considérations ont été développés [TH94]^[10] [WLB95]^[11] [Lam96]^[12] [Fra98]^[13].

Ces modèles sont complets dans le sens où ils intègrent les propriétés les plus pertinentes du SVH. Cependant, ils présentent dans leur ensemble un niveau de complexité et de temps de calcul qui freinent quelque peu leur utilisation notamment dans les applications temps réel.

model for target detection and recognition, Eli Peli, Editor, World Scientific, pp. 245-283, 1995.

- [5] M. Bolin and G. Meyer. Image quality assessment by using neural networks. *Special Interest Group on Computer Graphics and Interactive Techniques Conference, Vol. 5*, pp. 409-418, 1998.
- [6] M. Bolin and G. Meyer. A visual difference metric for realistic image synthesis. *SPIE, Human Vision and Electronic Imaging IV, Vol. 3644*, pp. 106-120, 1999.
- [7] J. Lubin. A human vision system model for objective picture quality measurements. *International Broadcasting Convention, Vol. 5*, pp. 498-503, 1997.
- [8] J. Lubin. Sarnoff JND vision model: Algorithm description and testing. *Technical report Sarnoff Corporation, Vol. 2*, pp. 485-488, 1997.
- [9] D.J. Heeger and P.C. Teo. A model of perceptual image fidelity. Heteroscedastic hough transform (htht): An efficient method for robust *International Conference on Image Processing, Vol. 2*, 1995.
- [10] P. C. Teo and D. J. Heeger. Perceptual image distortion. *Human Visual Processing and Digital Display, Vol. 2179*, pp. 127-141, 1994.
- [11] S. J. P. Westen, R. L. Lagendijk and J. Biemond. Perceptual image quality based on a multiple channel HVS model. *International Conference on Acoustics, Speech and Signal Processing*, pp. 2351-2354, 1995.
- [12] C. J. Van Den Branden Lambrecht. Perceptual Models and Architectures for Video Coding Applications. *PhD thesis, Ecole polytechnique Fédérale de Laussance*, 1996.
- [13] P. Franti. Blockwise distortion measure for statistical and structural errors in digital images. *Signal Processing: Image Communication, Vol. 13*, pp. 89-98, 1998.

1.4.4 Approches neuronales

L'intérêt des réseaux de neurones pour l'évaluation de la qualité d'image est tout à fait naturel. Il y a dans le processus d'évaluation subjective une forte dimension " apprentissage ". En effet, l'homme reconnaît, par exemple, facilement un flou sans avoir à sa disposition l'image originale.

Différentes méthodes d'estimation de la qualité sont basées sur des réseaux de neurones. Carrai et al. proposent une mesure de qualité basée sur l'extraction de caractéristiques locales de l'image originale et dégradée [CHGZ02]^[1]. Pour chaque bloc de l'image originale et de sa version dégradée, des descripteurs sont extraits à partir de l'histogramme (moyenne, écart-type, kurtosis, ...) et de la matrice de co-occurrence (entropie, énergie,...). Un réseau de neurones est ensuite utilisé pour estimer la qualité du bloc à partir de ces descripteurs. Une version améliorée a aussi été proposée dans [GZVH02]^[2]. Le même principe a été repris dans [BHB04]^[3]. Les descripteurs sélectionnés sont ici la moyenne et l'écart-type de l'image originale et de l'image dégradée ainsi que la covariance et l'erreur quadratique moyenne entre les deux images. Dans [DXH07]^[4], les descripteurs sont extraits dans le domaine des ondelettes. Chen et al. proposent quant à eux d'extraire les descripteurs à partir de l'image erreur (différence entre l'image originale et sa version dégradée) [CHG00]^[5].

Dans [DTZY08]^[6], les auteurs proposent d'utiliser les métriques SSIM et PSNR

-
- [1] P. Carrai, I. Heynderickx, P. Gastaldo P. and R. Zunino. Image quality assessment by using neural networks. *IEEE International Symposium on Circuits and Systems*, Vol. 5, pp. 253-256, 2002.
 - [2] P. Gastaldo, R. Zunino, E. Vicario and I. Heynderickx I. CBP neural network for objective assessment of image quality. *Conference on Proceedings of the International Joint*, Vol. 1, pp. 194-199, 2002.
 - [3] A. Bouzerdoum, A. Havstad and A. Beghdadi. Image quality assessment using a neural network approach. *IEEE International Symposium on Proceedings*, pp. 330-333, 2004.
 - [4] Y. Dongxue, H. Xinsheng and T. Hongli. Image Quality Assessment Based on Wavelet Coefficients Using Neural Network. *International Symposium on Neural Networks*, Vol. 2, pp. 853-859, 2007.
 - [5] S. Chen, Z. He and P.M. Grant. Artificial neural network visual model for image quality enhancement *Neurocomputing*, Vol. 30, pp. 339-346, 2000.
 - [6] W. Ding, Y. Tong, Q. Zhang and D. Yang. Image and Video Quality Assessment Using

comme entrée d'un réseau de neurones. Une machine à vecteur de support (SVM : Support Vector Machine) permet ensuite de discriminer entre les points aberrants de l'image. Une approche similaire basée sur SSIM et sa version dans le domaine ondelettes a été proposée dans [YWL05]^[7]. Une analyse en composantes principales sert d'étape de traitement intermédiaire entre les descripteurs (SSIM et SSIM ondelettes) et le réseau de neurones. Dans [PCD06]^[8], la méthode proposée combine une mesure d'estimation des effets de bloc [WBE00]^[9] ainsi que des descripteurs issus de l'histogramme 2D des gradients (amplitude et orientation) et d'une analyse temporelle entre deux images successives.

Ainsi, cet état de l'art nous a permis de voir combien il est difficile aujourd'hui à travers toutes les études menées ces 20 dernières années de faire le point et de pouvoir catégoriser le nombre considérable de méthodes d'évaluation de qualité d'image. Notre seul critère aujourd'hui est la corrélation avec les résultats des tests psycho-visuels. Mais là encore, l'absence de base de données acceptée de tous et de modèles de perception visuelle capable d'intégrer de façon fiable toutes les connaissances que nous avons du SVH rend la tâche difficile.

1.5 Conclusion

Ce chapitre a introduit en premier lieu la notion de qualité et de fidélité d'images. L'évaluation subjective a ensuite été présentée. Les différents protocoles y ont été décrits ainsi que les principaux facteurs influençant l'appréciation humaine. Le MOS (Mean Opinion Score) obtenu à partir de la compilation des notes subjectives a ensuite été défini. A titre d'exemple, des bases d'images ont

Neural Network and SVM. *Tsinghua Science Technology*, Vol. 13, pp. 112-116, 2008.

- [7] Y. You, A. Wang and F. Lu. A Wavelet Approach To Image Quality Assessment Using Neural Networks. *Electrical and Computer Engineering, University of Texas at Austin*, Vol. 17, pp. 1316-1327, 2005.
- [8] P. Le Callet, C. Viard-Gaudin and D. Barba. A convolution neural network approach for objective video quality assessment. *IEEE Transactions on Neural Networks*, Vol. 17, pp. 1316-1327, 2006.
- [9] Z. Wang, A.C. Bovik and B.L. Evans. Blind measurement of blocking artefacts in images. *International Conference on Image Processing*, 2000.

été brièvement décrites. Cette première partie permet de constater la difficulté de mise en place de tels tests. De ce fait, les mesures objectives sont vues comme une solution alternative.

Ainsi, la notion d'évaluation objective a été introduite. Nous avons dans un premier temps défini les différentes catégories de mesures de qualité, à savoir les mesures avec référence, sans référence et avec référence réduite, en soulignant leurs différences et leurs apports. Des mesures classiques basées sur des considérations orientées purement signal ont été présentées. Une description sommaire du fonctionnement du SVH et des modèles associés, nous a permis d'introduire les métriques inspirées essentiellement par le SVH. Des mesures basées sur une analyse neuronale ont aussi été présentées. La figure 1.24 récapitule l'ensemble des notions abordées dans ce chapitre ainsi que celles qui seront évoquées dans les chapitres suivants.

La suite de ce rapport de thèse est composée de deux parties, la première est consacrée à l'étude des dégradations dues à la compression irréversible, et plus particulièrement les effets de bloc et le ringing. Nous nous intéressons aussi au flou qui peut provenir également de la compression ou d'autres origines. L'étude de ces dégradations est faite sans référence. Un chapitre est ainsi dédié à chacune de ces distorsions et débute par une présentation de leurs origines. Un bref état de l'art et les solutions proposées dans ce travail y sont ensuite présentés.

La deuxième partie, dédiée à la proposition d'un schéma universel d'estimation de qualité d'image avec et sans référence, se compose de deux chapitres à travers lesquels nous mettons en évidence l'inconvénient majeur de ces méthodes. Dans le premier chapitre, nous proposons un schéma de fusion permettant d'améliorer les performances des métriques actuelles. Les solutions proposées y seront détaillées ainsi que les résultats obtenus. Dans le deuxième chapitre, le système de mesure complet est défini et décrit en détails.

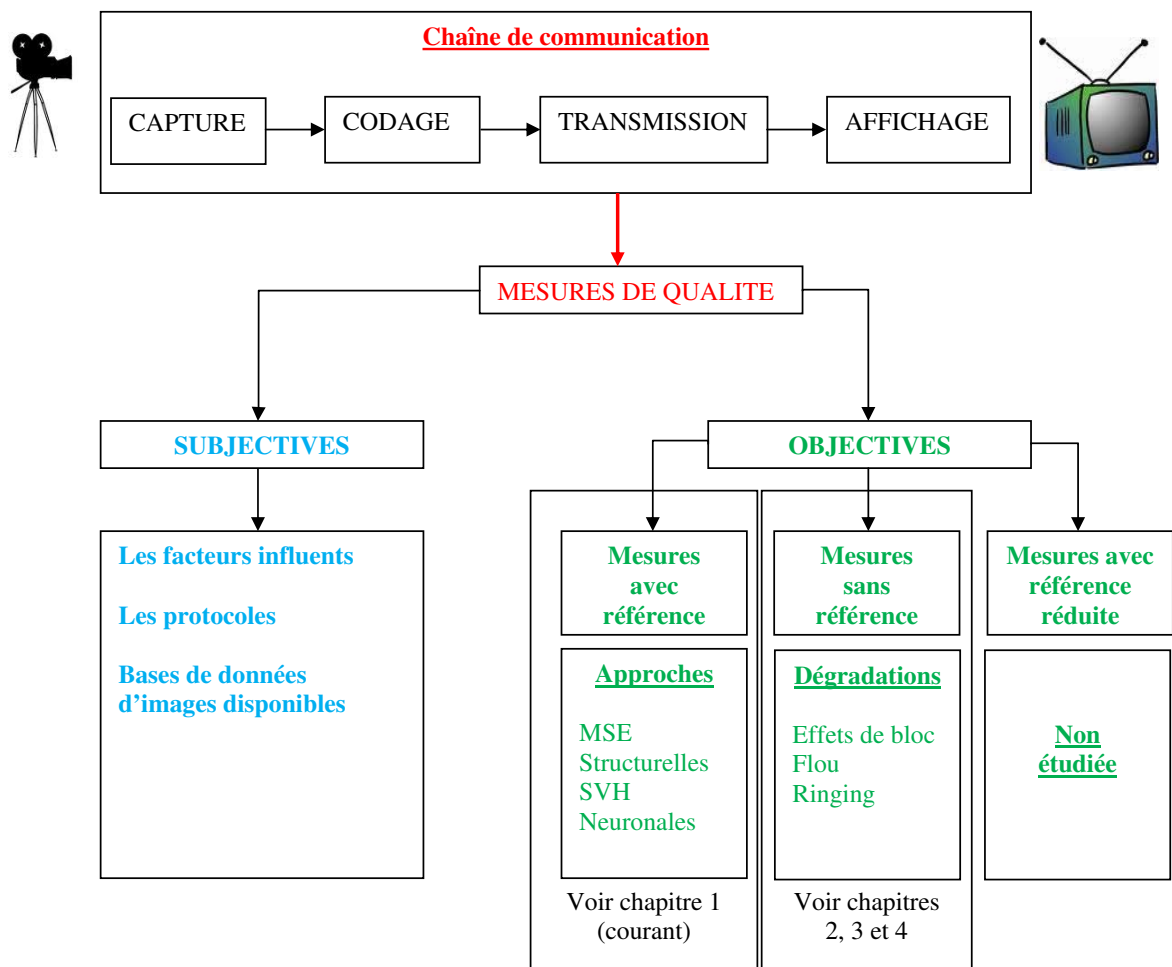


FIGURE 1.24 – Schéma récapitulatif.

Première partie

Etudes des dégradations

Chapitre 2

Etudes des dégradations de compression par bloc

Sommaire

2.1	Introduction	66
2.2	Etat de l'art	67
2.2.1	Approches spatiales	67
2.2.2	Approches fréquentielles	69
2.3	Prédiction des effets de blocs	70
2.3.1	Description de la méthode	71
2.3.2	Résultats expérimentaux	76
2.3.3	Application : Tatouage	78
2.4	Estimation de la visibilité des effets de bloc	83
2.4.1	Carte de visibilité	83
2.4.2	Réduction des effets de bloc	88
2.5	Conclusions	93

2.1 Introduction

Dans ce chapitre, nous étudions une des dégradations les plus répandues dans le signal image, à savoir les effets de bloc. Visuellement, cette distorsion se manifeste généralement au niveau des frontières entre blocs et apparaît comme des contours verticaux et horizontaux dont la visibilité dépend fortement de la distribution spatiale du signal image. La régularité de ces contours occasionne une gêne importante et rend alors cet artefact très visible. De plus, la sensibilité accrue du SVH aux contrastes verticaux et horizontaux, et l'accroissement de la perception des contrastes entre deux régions adjacentes ("effet de Mach" [Wel68]^[1]) ne font qu'accroître la gêne occasionnée par cet artefact. La figure 2.1 illustre un exemple d'images dégradées.

Les effets de bloc sont la conséquence d'un découpage de l'image en blocs et leur traitement de façon indépendante. C'est généralement le cas des méthodes de compression par bloc telles que JPEG ou la quantification vectorielle.

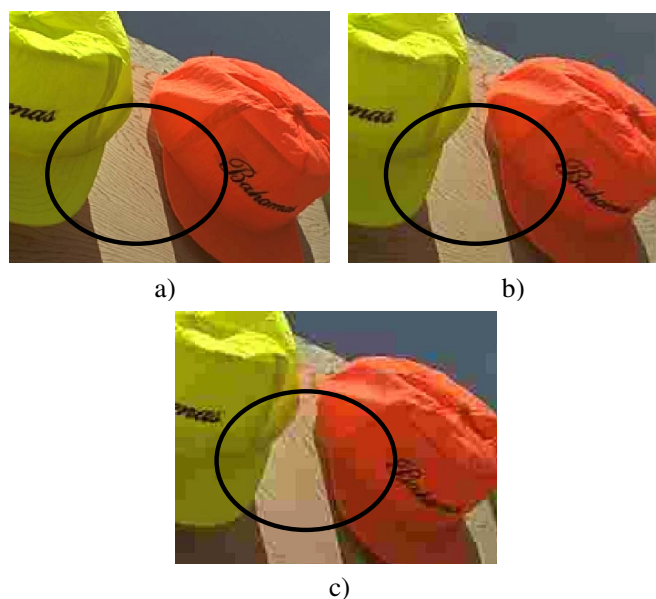


FIGURE 2.1 – a) Image originale et ses b-c) versions compressées .

Dans ce qui suit, nous présentons d'abord un bref état de l'art des métriques

[1] W.T. Welford. The visual Mach effect. *Physics Education*, Vol. 3, pp. 83-88, 1968.

d'estimation des effets de bloc sans référence. Nous décrivons ensuite les études liées respectivement à la prédiction des effets de bloc et à leur estimation.

2.2 Etat de l'art

L'effet de bloc est probablement la dégradation la plus fréquente et la plus étudiée. Plusieurs méthodes sans référence ont été proposées pour estimer cette distorsion. Ces méthodes peuvent être classées en deux catégories : les mesures spatiales, qui se concentrent sur les caractéristiques spatiales des effets de bloc, et les métriques fréquentielles qui exploitent ses singularités dans le domaine transformé. Nous présentons ci-après un état de l'art des mesures développées dans ces deux domaines.

2.2.1 Approches spatiales

La méthode proposée par Coudoux et al. [CGC98]^[2] est un premier exemple de méthode d'estimation des effets de bloc par une analyse spatiale. Le contraste et la déformation non linéaire de l'écran (correction gamma) sont intégrés et utilisés pour estimer l'impact visuel de la dégradation.

L'idée a été reprise et développée par Wang et al. [WSB02]^[3]. Ces auteurs proposent en plus de l'analyse du contenu spatial de chaque bloc (intra-bloc), la variation des gradients. Les valeurs obtenues sont ensuite pondérées pour en déduire une mesure globale.

Pan et al. proposent d'estimer la visibilité des frontières d'un bloc donné (A) en fonction du bloc adjacent supérieur (B) et celui de gauche (C) [PLR⁺04]^[4]. Ainsi,

-
- [2] F.X. Coudoux, M.G. Gazelet and P. Corlay. Reduction of blocking effect in DCT-coded images based on a visual perception criterion. *Signal Processing: Image Communication*, Vol. 11, pp. 179-186, 1998.
 - [3] Z. Wang, H. Sheikh and A. Bovik. No-reference perceptual quality assessment of JPEG compressed images. *International Conference on Image Processing*, pages 477-480, 2002.
 - [4] F. Pan, X. Lin, S. Rahardja, W. Lin, E. Ong, S. Yao, Z. Lu and X. Yang. A locally-adaptive algorithm for measuring blocking artifacts in images and videos. *International Symposium*

la visibilité de la frontière supérieure (ou gauche) est déterminée par l'analyse inter-bloc entre le bloc A et B (ou C). Ces visibilitées sont ensuite comparées à un seuil JND (Just Noticeable Distortion) et combinées pour former une mesure de qualité.

Dans [BCFM02]^[1], Bailey et al. reprennent le même principe en intégrant au préalable une étape de détection des effets de bloc. Cette détection est réalisée par projection des gradients verticaux et horizontaux de l'image. Cette opération laisse apparaître un motif régulier (succession de pics) non présent généralement dans les images naturelles non dégradées. Ces pics révèlent ainsi la présence ou non de la distorsion. L'idée a aussi été exploitée dans [MK05]^[2].

Une approche différente a été proposée par Babu et al. [BP05]^[3]. Les auteurs proposent d'estimer la qualité d'une image à travers l'extraction de descripteurs locaux. Une fois extraits, ils sont ensuite utilisés à l'entrée d'un réseau de neurones qui fournit en sortie une estimation de la qualité globale de l'image.

Dans [SY09]^[4], Song et al. proposent d'estimer la qualité d'une image en intégrant la notion de masquage. La visibilité des effets de blocs est supposée inversement proportionnelle à la luminosité du fond (effet de masquage). Le masquage de textures appelé par les auteurs "complexity masking" a aussi été intégré.

Crête et al. introduisent la notion de "bloc étendu" [Crê07]^[5]. Cette notion fait référence à la disparition de certaines (fausses) frontières dans l'image et à

-
- on Circuits and Systems, Vol. 3, pp. 925-928, 2004.*
- [1] D. Bailey, M. Carli, M. Farias and S. Mitra. Quality assessment for block-based compressed images and videos with regard to blockiness artifacts. *International Workshop in Data Compression*, 2002.
 - [2] R. Muijs and I. Kirenko. A No-Reference Blocking Artifact Measure for Adaptive Video Processing. *European Signal Processing Conference*, 2005.
 - [3] R.V. Babu and A. Perki. An hvs-based no reference perceptual quality assessment of jpeg coded images using neural networks. *International Conference on Image Processing*, pp. 433-439, 2005.
 - [4] X. Song and Y. Yang. A new no reference assessment metric of blocking artefacts based on HVS masking effect. *International Congress on Image and Signal Processing*, pp. 1-6, 2009.
 - [5] F. Crête. Estimer, mesurer et corriger les artefacts de compression pour la télévision. *Rapport de Thèse, Université Joseph Fourier*, 2007.

l'apparition ou la fusion de blocs homogènes. Ce phénomène est illustré par la figure 2.1. Les auteurs proposent d'intégrer cette notion en classant d'abord le type de bloc dans l'image (bloc étendu ou non). Cette classification est réalisée par l'utilisation de plusieurs seuils empiriques. Un poids de visibilité est ensuite attribué à chaque frontière dans l'image.

2.2.2 Approches fréquentielles

Notons que les effets de blocs peuvent être aussi estimés dans le domaine transformé. Cependant, peu de métriques ont été proposées. On peut citer la méthode proposée par Wang et al. [WBE00]^[6] qui est basée sur l'analyse du spectre de Fourier. Le principe est de détecter les pics qui correspondent aux fréquences de découpage de l'image en blocs 8x8. L'image des gradients verticaux et horizontaux est d'abord calculée et transformée en signaux 1D. Puis, le spectre de chaque segment est calculé et comparé à sa version filtrée (filtre médian). Cette mesure est ensuite utilisée pour estimer la distorsion. Dans [OKD09]^[7], la métrique a été reprise et utilisée pour estimer la qualité des vidéos.

Bovik et al. proposent d'estimer les effets de bloc dans le domaine TCD [BL01]^[8]. Les effets de bloc sont modélisés comme une fonction 2D à partir de la distribution de l'information entre deux blocs adjacents. Les coefficients TCD de ce bloc sont d'abord calculés et analysés. La visibilité d'un bloc est finalement dérivée de cette analyse.

Il existe d'autres méthodes d'estimation des effets de blocs [LLP98]^[9] [Zen99]^[10]. Il est à noter que l'ensemble des méthodes proposées sont basées sur une in-

-
- [6] Z. Wang, A.C. Bovik and B.L. Evans. Blind measurement of blocking artefacts in images. *International Conference on Image Processing*, 2000.
 - [7] T. Oelbaum, C. Keimel and K. Diepold. Rule based no reference video quality evaluation using additionally coded videos. *IEEE Journal of Selected Topics in Signal Processing*, Vol. 3, pp. 294-303, 2009.
 - [8] A.C. Bovik and S. Liu. DCT-domain blind measurement of blocking artefacts in DCT-coded images. *International Conference on Acoustics, Speech, and Signal Processing*, 2001.
 - [9] Y.L. Lee, H.C. Lim and H.W. Park. Blocking effect reduction of JPEG images by signal adaptative filtering. *IEEE Transactions Image Processing*, 1998.
 - [10] B. Zeng. Reduction of blocking effect in DCT-coded images using zero-masking techniques.

formation a priori pas toujours vérifiée. En effet, les mesures présentées sont généralement centrées sur la taille du bloc. Cette information est supposée connue.

Dans le cadre de cette étude, nous proposons une approche qui s'affranchit de cette restriction en considérant les effets de bloc comme une dégradation qui tend à uniformiser certaines régions de l'image.

Dans la suite, nous présentons une méthode de prédiction des effets de bloc basée sur un apprentissage. Puis, nous décrivons une deuxième méthode dont l'objectif est d'estimer la distorsion à partir de l'image dégradée.

2.3 Prédiction des effets de blocs

Nous abordons ici une problématique différente de celle généralement étudiée dans la littérature. Nous proposons de prédire l'apparition des effets de blocs avant compression. L'intérêt d'une telle opération est d'offrir le moyen d'ajuster les paramètres de codage de façon à améliorer la qualité de l'image.

Plus précisément, pour une méthode de compression par bloc donnée et un taux de compression donné, nous associons à chaque région de l'image originale (avant compression), un poids représentant la visibilité (probable) des effets de blocs après compression. Ce poids peut être interprété comme étant une probabilité d'apparition des effets de bloc. Il s'agit en d'autres termes de proposer un élément décisif qui peut être intégré dans un schéma de compression avec contrôle de qualité.

Dans le cadre de cette étude, nous nous focalisons sur les méthodes de compression par bloc et plus particulièrement sur celles basées sur la TCD, telles que JPEG et MPEG2. Il est à noter que la méthode peut être étendue à l'ensemble des méthodes de compression par bloc telles que JPEG XR, H.264.

2.3.1 Description de la méthode

L'idée principale développée ici est d'étudier la relation entre l'apparition des effets de bloc et le voisinage des pixels dans l'image non compressée. On en déduira alors une fonction permettant d'affecter à chaque pixel de l'image non compressée un poids, pouvant être interprété comme la probabilité (locale) d'apparition d'effets de bloc. Pour ce faire, nous effectuons un apprentissage sur une base d'images en niveau de gris. Le schéma synoptique est illustré par la figure 2.2.

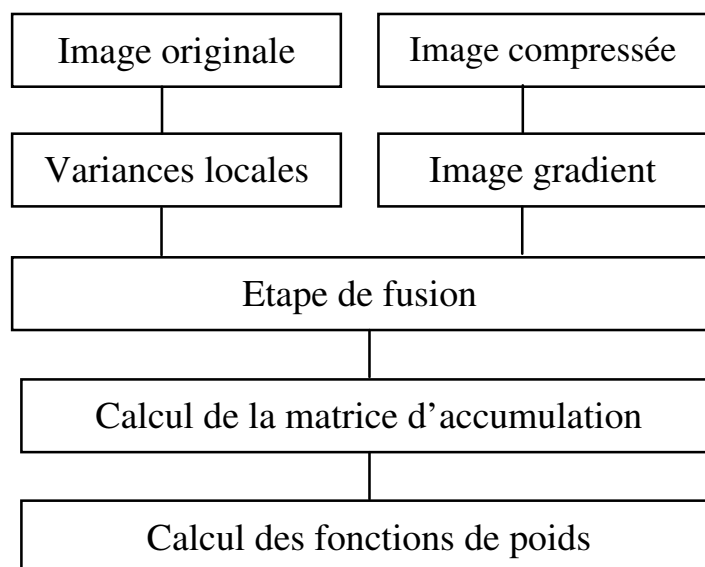


FIGURE 2.2 – Schéma synoptique.

L'apprentissage est réalisé sur une base de plus de 200 images naturelles (F. C. Donders Centre for Cognitive Neuroimaging database, voir figure 2.3) dont le contenu est variable (différents types de textures, de distributions spatiales et de contrastes).

Avant d'extraire certaines caractéristiques locales des images de la base d'apprentissage, nous analysons, dans un premier temps, la distribution spatiale des pixels. En effet, l'apparition des effets de bloc dans l'image dépend fortement de descripteurs locaux comme la couleur, l'homogénéité, le gradient, etc. Ces caractéristiques locales sont souvent exprimées dans l'espace des fréquences spatiales



FIGURE 2.3 – Échantillon d'images de la base d'apprentissage.

par des transformées telles que la transformée en ondelettes ou la TCD. Cependant, afin de rester le plus indépendant possible de la méthode de compression, nous choisissons comme descripteur la variance locale du signal dans le domaine spatial. Celle-ci peut être utilisée comme une mesure statistique simple permettant de décrire l'homogénéité et l'activité locale du signal. En effet, l'étape de quantification génère de larges zones uniformes pouvant être identifiées et détectées en analysant l'uniformité du signal image (gradient faible ou nul).

Ainsi, pour chaque image \mathbf{f} de la base d'apprentissage, on calcule l'image des variances locales correspondante, notée ici \mathbf{V} (voir figure 2.4.b). La variance locale est ici calculée sur une fenêtre de taille 3×3 .

On analyse alors statistiquement la relation entre ces variances locales et l'apparition des effets de bloc dans les images compressées à différents taux de compression.

Avant d'aller plus loin dans la description de la méthode proposée, nous introduisons les notations suivantes :

- g_q : L'image compressée \mathbf{f} , \mathbf{q} représente les différents facteurs de qualité (\mathbf{q} [1,100] pour JPEG).
- g'_q : L'image des gradients correspondante.

Ce premier processus ne permet d'obtenir qu'une détection simple et grossière

des effets de bloc (voir figure 2.4.c). On montrera, par la suite, que les erreurs induites par cette détection restent négligeables et sans effet sur le processus d'apprentissage.

On utilise ici une méthode d'apprentissage cumulative basée sur un processus de vote. On définit alors une table d'accumulation qui représente les statistiques d'apparition des effets de bloc dans les images compressées pour différents facteurs de qualité. Cette table de vote est une matrice 2-D notée $H(v, q)$ où v correspond à la valeur de la variance et q représente le facteur de qualité ($q \in [1, 100]$). Il est à noter que pour pouvoir appliquer la méthode à d'autres techniques de compression, il suffit d'appliquer le même principe en modifiant uniquement les taux de compression.

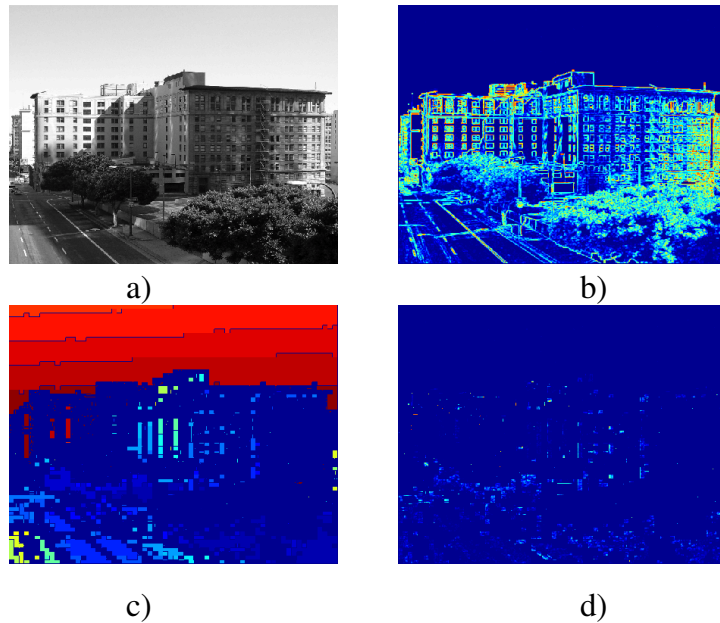


FIGURE 2.4 – a) Image originale, b) Image des variances locales, c) Image des gradients nuls et d) Image des gradients pondérés par l'image des variances locales.

Notons $\{f^i, i \in [1, N]\}$ l'ensemble des images de la base d'apprentissage et V^i , g_q^i , $g_q'^i$ les images des variances, compressées et gradients respectivement.

Soit (x, y) un pixel de l'image f^i . Pour chaque couple (v, q) , on définit la fonction d'influence du pixel (x, y) par :

$$\varphi^i(v, q) = \begin{cases} T(x^i, y^i) & \text{si } v = V^i(x^i, y^i) \text{ et } g_q'^i(x^i, y^i) = 0 \\ 0 & \text{sinon} \end{cases} \quad (2.1)$$

avec $T(x^i, y^i)$ la visibilité du pixel (\mathbf{x}, \mathbf{y}) de l'image \mathbf{i} .

La carte de visibilité est obtenue comme suit :

- Une première phase consiste à filtrer l'image originale et sa version dégradée dans le domaine de Fourier. Cette opération permet de tenir compte de la sensibilité du SVH à détecter certains détails de l'image. Le filtre CSF de Daly est ici utilisé [Dal93]^[1].
- Une décomposition multi-canal (transformée en Cortex) est ensuite appliquée aux deux images. Cette décomposition permet de reproduire la sélectivité fréquentielle et directionnelle du SVH.
- Pour chaque sous-bande, un modèle de masquage est appliqué. La visibilité de la dégradation est ainsi estimée. Le modèle de masquage appliqué est celui proposé par Daly [Dal93]^[1].
- La carte de visibilité de la dégradation est finalement obtenue par combinaisons des cartes de visibilité obtenues pour chaque sous bande. Cette combinaison est ici réalisée par sommation des différentes cartes de visibilité.

L'équation 2.1 signifie qu'un pixel (x^i, y^i) peut avoir une influence dans la cellule (\mathbf{v}, \mathbf{q}) de $\mathbf{H}(\mathbf{v}, \mathbf{q})$ si sa variance locale correspond à \mathbf{v} et que la valeur absolue de son gradient ($g_q'^i$) dans l'image compressée est nulle (le pixel est très probablement affecté par la dégradation).

Après avoir calculé les fonctions d'influence pour les pixels (\mathbf{x}, \mathbf{y}) des images f^i de la base d'apprentissage, on définit les valeurs d'une cellule (\mathbf{v}, \mathbf{q}) de la table d'accumulation \mathbf{H} comme suit :

[1] S. Daly. The visible differences predictor: an algorithm for the assessment of image fidelity. *Digital images and human vision*, MIT Press, 1993.

$$H(v, q) = \sum_i \varphi^i(v, q) \quad (2.2)$$

La figure 2.5 montre un exemple de table d'accumulation. Pour une meilleure visibilité, uniquement les variances entre 0 et 16 y sont représentées. Cette figure montre clairement la cohérence des données relevées. En effet, on peut constater que plus la variance et le facteur de qualité sont faibles, plus la probabilité d'apparition des effets de blocs est élevée. Les erreurs dues à l'approximation de la détection des effets de bloc ont donc complètement disparu grâce à la masse de données correctement accumulées.

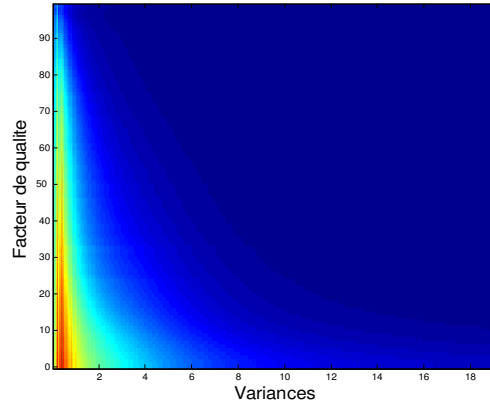


FIGURE 2.5 – Matrice d'accumulation des votes.

A partir de la matrice de votes, une fonction de poids représentant la probabilité locale d'apparition des effets de bloc (après compression) selon le voisinage des pixels (caractérisés par leurs variances) dans l'image non compressée est calculée. Pour un facteur de qualité donné, la fonction de poids peut être simplement obtenue en exploitant la matrice \mathbf{H} :

$$w_q(v) = H(v, q) \text{ avec } v \in [0, 255] \quad (2.3)$$

La matrice \mathbf{H} est obtenue à partir de la base d'apprentissage contenant des images réelles. Il est important de souligner que toutes les valeurs possibles des variances locales ne sont pas forcément représentées. De plus, les régions homogènes

sont prédominantes dans les images naturelles (les valeurs de variances faibles sont plus représentées que les variances élevées). Ceci conduit à une sous représentation de certaines cellules de la table de votes. Pour pallier ce problème, on utilise une interpolation polynômiale de la fonction de poids (équation 2.3). Après normalisation, on obtient les fonctions de poids illustrées par la figure 2.6 (pour $q = 80, 25, 15$ et 10).

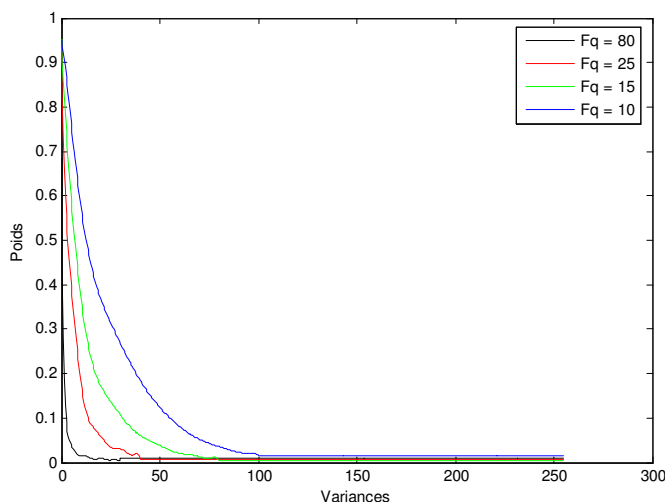


FIGURE 2.6 – Fonction de poids pour les facteurs de qualité de 80, 25, 15 et 10.

On peut là aussi noter la cohérence des résultats obtenus. En effet, pour une variance donnée v , plus le facteur de qualité diminue (bas débit), plus le poids est élevé (la probabilité d'apparition des effets de bloc augmente).

2.3.2 Résultats expérimentaux

Pour évaluer les performances de la méthode, des tests expérimentaux sont réalisés à différents taux de compression sur une base de plus de 200 images naturelles différentes de celles utilisées durant l'apprentissage. La procédure expérimentale est très simple et ne requiert que l'image originale. Étant donné une image test f à analyser, la probabilité pour qu'un pixel (x, y) appartienne à une région affectée par les effets de bloc, pour un facteur de qualité q de compression (JPEG) est donnée par :

$$w_q(v) = H(v, q) \text{ avec } v = V(x, y) \quad (2.4)$$

où w_q représente la fonction de poids obtenue par apprentissage et \mathbf{V} la variance locale en (\mathbf{x}, \mathbf{y}) .

Les fonctions de poids (obtenues durant la phase d'apprentissage) peuvent être vues comme de simples tables de correspondances (LUT). La procédure de prédiction est donc réduite à un simple calcul de variances locales.

La figure 2.7 représente un exemple de prédiction des effets de blocs à différents facteurs de qualité où les régions rouges et bleues correspondent respectivement aux fortes et faibles probabilités d'apparition de ces artefacts.

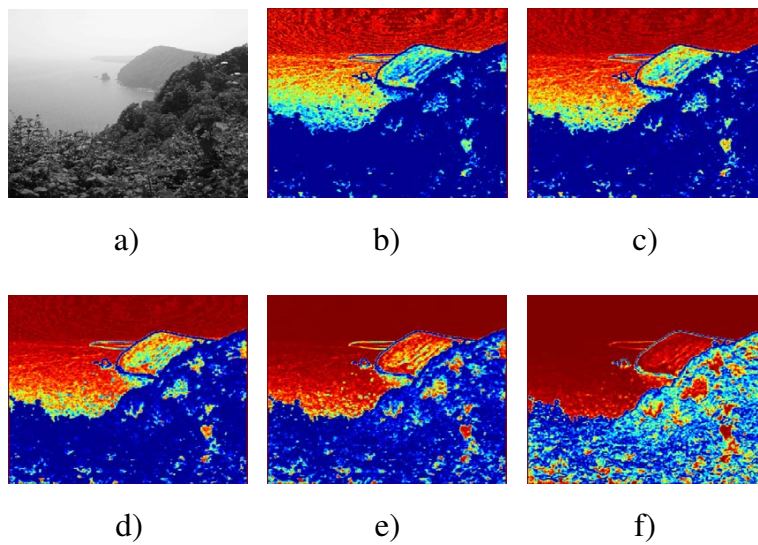


FIGURE 2.7 – a) Image originale b-f) Résultats obtenus pour $Fq=30$, $Fq=20$, $Fq=15$, $Fq=10$ et $Fq=5$, respectivement.

À travers ces résultats, on voit bien que les probabilités augmentent lorsque le facteur de qualité diminue. De plus, il apparaît que les régions homogènes sont plus affectées par les effets de blocs (avec la diminution de q) que les zones texturées. Les mêmes constatations ont été faites pour différentes images tests.

Un autre exemple est illustré par la figure 2.8. L'image originale, compressée

ainsi que la carte de prédiction y sont présentées. On constate que les poids forts correspondent bien aux régions les plus dégradées par les artefacts de compression.

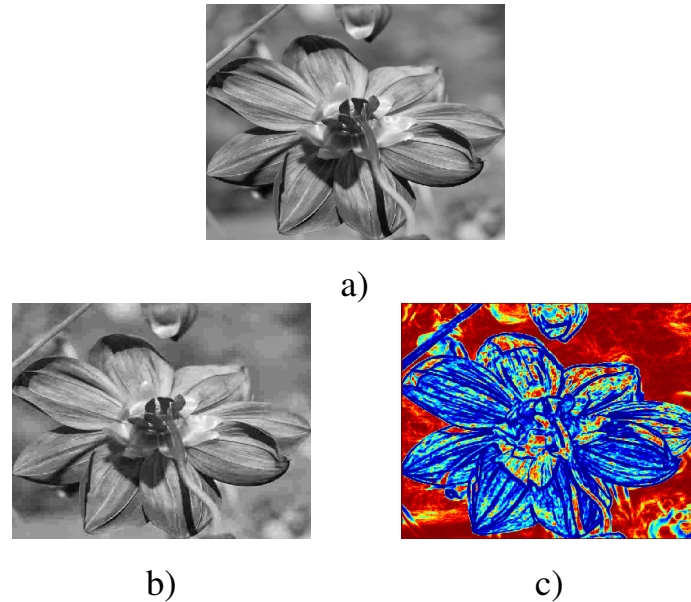


FIGURE 2.8 – Résultat obtenu sur une image naturelle.

2.3.3 Application : Tatouage

Dans cette section, nous proposons d'utiliser la carte de prédiction des effets de bloc pour augmenter la robustesse et améliorer la qualité d'un schéma tatouage donné face aux attaques de type compression par bloc. En effet, cette carte permet d'identifier les régions de l'image les plus vulnérables aux effets de la compression par bloc telle que JPEG. En exploitant cette carte, on peut montrer qu'il est possible d'atteindre un niveau de robustesse assez intéressant [CMB08]^[1].

Dans la section suivante, nous présentons la stratégie de tatouage proposée et discuterons ses performances.

Description de la méthode :

-
- [1] A. Chetouani, G. Mostafaoui and A. Beghdadi. Predicting blocking effects in the spatial domain using a learning approach. *International Conference on Signal Processing and Multimedia Applications*, 2008.

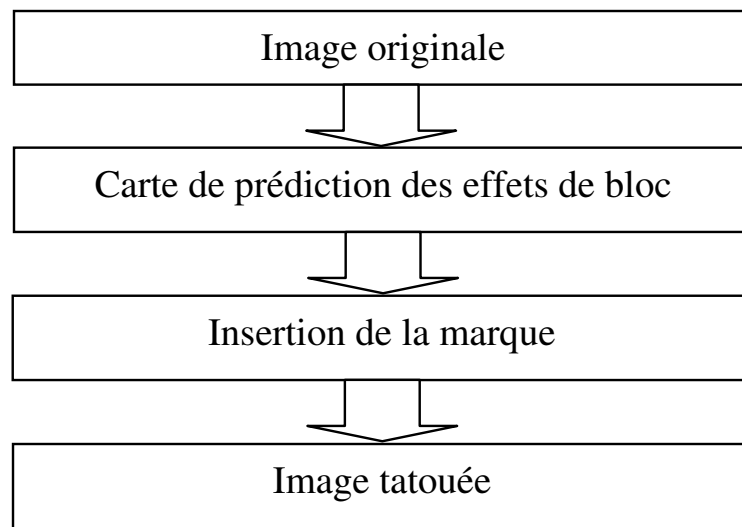


FIGURE 2.9 – Schéma synoptique.

La carte de prédiction des effets de bloc (BEP) est d'abord binarisée en utilisant un seuillage empirique (BBEP : Binary BEP). Cette opération nous permet de déterminer les zones de l'image les moins vulnérables aux effets de compression (les zones pertinentes ou saillantes de l'image, voir figure 2.10). Il est à noter qu'un seuil égal à 0.5 ($Th_M = 0.5$) signifie que la probabilité qu'un pixel donné soit affecté par la compression est de 50%. Après une campagne de tests sur les différentes images, il a été constaté qu'en fixant le seuil à 0.3, on réalise un bon compromis entre transparence et robustesse dans le schéma de tatouage.

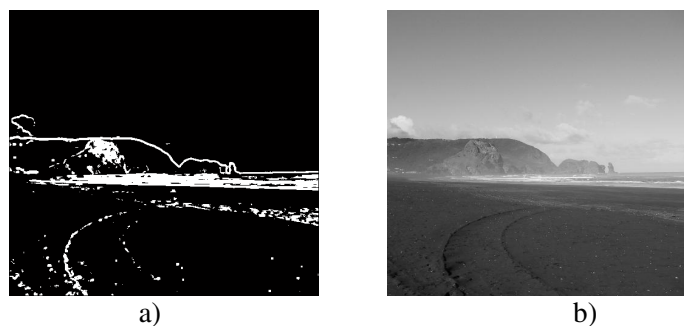


FIGURE 2.10 – a) Version binarisée de la carte de prédiction de b) l'image "Beach".

La carte binaire est ensuite analysée pour repérer les régions les plus stratégiques pour insérer le tatouage. Un bloc est considéré comme zone d'insertion robuste si la proportion de pixel à 1 est supérieure à 0.5 (ici 32/64).

La marque est alors ensuite insérée dans les blocs ainsi sélectionnés. Ici, une méthode de tatouage simple est utilisée. Elle est basée sur l'insertion d'une séquence bipolaire $(-/+1)$ pseudo-aléatoire de moyenne nulle et de variance unitaire dans le domaine TCD. La marque est insérée uniquement dans une certaine bande de fréquence. La loi d'insertion est définie par :

$$C_k(u, v) = C_k(u, v) + a.BBEP_k.V(u, v).W \quad (2.5)$$

où a décrit la force de tatouage, fixé à 15. $C_k(u, v)$ est le coefficient TCD à la fréquence (u, v) dans le bloc k de l'image. $V(u, v)$ correspond au masque de sélection fréquentiel (voir figure 2.11).

$$V(i, j) = \begin{cases} 1, & \text{si Bande de fréquence centrale} \\ 0, & \text{si non} \end{cases} \quad (2.6)$$

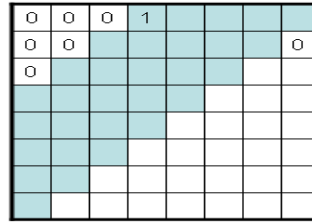


FIGURE 2.11 – Masque fréquentiel.

Le schéma de tatouage ainsi proposé est évalué en termes de transparence et de robustesse, face à différentes attaques.

Evaluation de transparence :

L'objectif principal de ce travail est d'améliorer la robustesse des méthodes de tatouage face aux attaques (volontaires ou involontaires) de type JPEG et ceci

Image	Without BEP Map		With BEP Map	
	PSNR	SSIM [WBSS04] ^[1]	PSNR	SSIM [WBSS04] ^[1]
Lena	27.8959	0.74016	30.2354	0.88545
Barbara	27.8973	0.79196	29.3513	0.89176
Peppers	27.9329	0.72577	30.0671	0.86601
Parrots	27.911	0.69504	31.9292	0.90406
Baboon	27.8986	0.86955	28.0063	0.87643
Man	28.2005	0.83868	28.5944	0.86545
Couple	27.9231	0.80942	28.5023	0.84717
Clown	27.8982	0.77799	28.8983	0.85581

TABLE 2.1 – Evaluation de la transparence.

en garantissant la qualité visuelle de l'image. L'évaluation est donc restreinte à la comparaison des résultats avec et sans utilisation de la carte BEP. La figure 2.12 présente un résultat de tatouage. On peut voir que l'utilisation de la carte BEP permet d'améliorer nettement la qualité visuelle de l'image et donc de réduire la visibilité des artefacts engendrés lors de l'insertion de la marque.



FIGURE 2.12 – a) Image tatouée avec la carte BEP et b) Image tatouée sans la carte BEP.

La transparence de la marque est en outre évaluée objectivement sur un ensemble de 8 images. Le tableau 2.1 présente les résultats obtenus. Une nette amélioration du degré de transparence, mesurée au moyen des métriques de qualité objectives, est obtenue par la méthode proposée.

Evaluation de la robustesse :

Q	5	10	15	20
Taux de réussite (%)	98.89	93.33	86.67	84.44

TABLE 2.2 – Pourcentage de réussite.

Pour évaluer les performances en termes de robustesse de la méthode proposée, la distribution du détecteur de la marque (sortie) obtenue avec et sans la carte BEP est présentée (figure 2.13). On s’aperçoit que la corrélation a considérablement augmentée (environ 3 fois plus grande). Ce qui signifie que la marque est relativement mieux détectée.

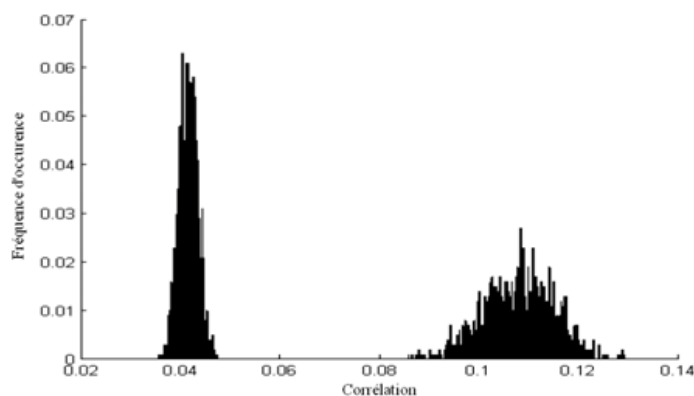


FIGURE 2.13 – Distribution du détecteur de sortie avec (droite) et sans (gauche) utilisation de la carte BEP pour l’image ”Beach” avec 1000 marques différentes.

D’autres expériences ont également été menées sur un ensemble de 90 images de la base CNN à différents facteurs de qualité. Le tableau 2.2 affiche le taux de réussite (pourcentage d’images dont une plus forte corrélation est obtenue par la méthode proposée). Les résultats obtenus montrent clairement l’augmentation des performances. On s’aperçoit aussi que le taux de réussite (TR) est toujours élevé pour un débit de compression moyen ($Q = 20$) et qu’il diminue lorsque celui-ci augmente. Cela est dû essentiellement au fait qu’à bas niveau de compression (Q élevé), très peu de zones sont affectées par la compression ; la marque est donc uniformément répartie sur l’ensemble de l’image. Par conséquent, la solution proposée tend vers le schéma d’insertion classique.

La robustesse de la méthode proposée contre diverses attaques a également été testée. Le tableau 2.3 présente les résultats obtenus pour l’image ”Beach”.

Attaques	Résultats
JPEG	Q=5
JPEG 2000	0.04bpp
Bruit gaussien	m=0, $\sigma = 40\%$
Zoom	50x50
Filtre Wiener	yes
Filtre moyennneur	3x3
Changement d'échelle	0.3
Egalisation d'histogramme	yes

TABLE 2.3 – Evaluation de la robustesse (image "beach").

Il est à noter que l'objectif ici est de proposer une application directe de la carte BEP.

2.4 Estimation de la visibilité des effets de bloc

Dans cette section, nous nous intéressons cette fois-ci à l'estimation de la dégradation (i.e. après compression). L'objectif ici est de calculer une carte de visibilité perceptuelle des effets de bloc. Cette carte est ensuite utilisée pour réduire les effets de la distorsion. Le principe de la méthode proposée est illustré par la figure 2.14.

2.4.1 Carte de visibilité

Les méthodes de compression par bloc ont tendance à créer des effets de bloc qui se manifestent par l'apparition de régions uniformes dans l'image. Nous proposons alors de calculer la carte de visibilité en exploitant cette caractéristique.

La première phase consiste à détecter les régions dégradées en analysant l'activité locale du signal image (gradient faible ou nul). Nous proposons d'identifier ces régions en segmentant d'abord l'image. À l'issue de cette segmentation, deux types de régions sont identifiées (C1 et C2). La première classe correspond aux régions les plus affectées par la distorsion (figure 2.15.b). Tandis que la seconde classe correspond aux régions qui ne seront pas ou peu altérées (figure 2.15.c).

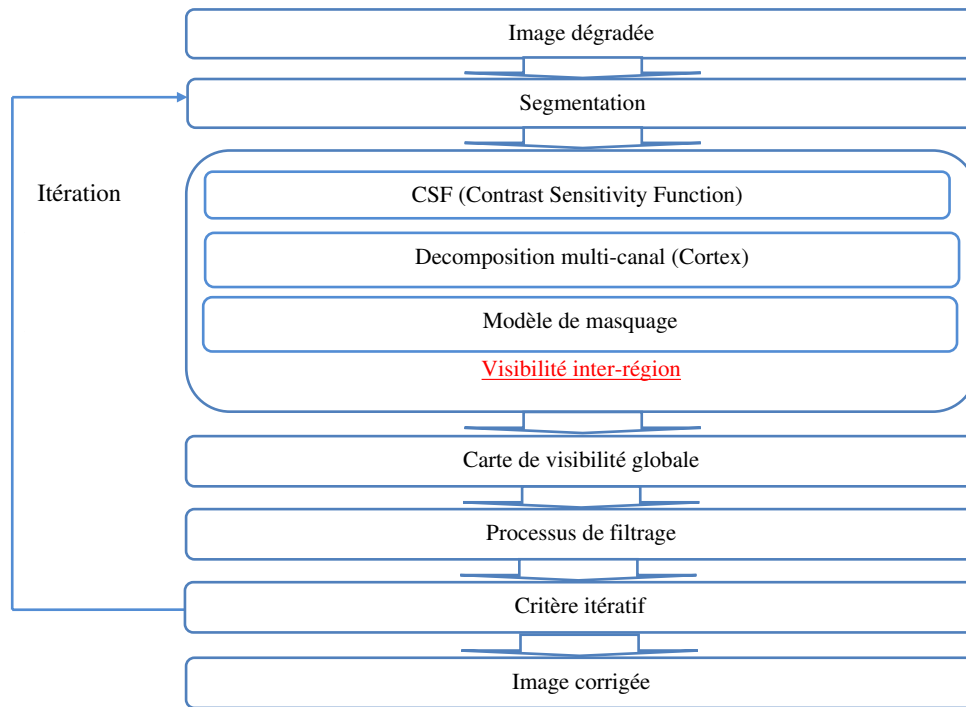


FIGURE 2.14 – Schéma synoptique.

Cette opération peut être réalisée de différentes manières, cependant pour rester le plus indépendant possible de la méthode de compression, nous choisissons comme descripteur la variance locale. Un algorithme d'étiquetage est ensuite appliqué afin de délimiter les régions.

Un résultat de segmentation est présenté figure 2.15.d. Nous introduisons les notations suivantes :

- R_i : $i^{\text{ème}}$ région de l'image segmentée.
- N_{ij} : $j^{\text{ème}}$ région adjacente de la région R_i .

Il est à noter que la visibilité des frontières d'une région donnée dépend fortement de ses régions adjacentes. En effet, une région peut être affectée par la dégradation sans pour autant être visible. Ainsi, nous proposons d'analyser l'image en tenant compte de ses caractéristiques locales.

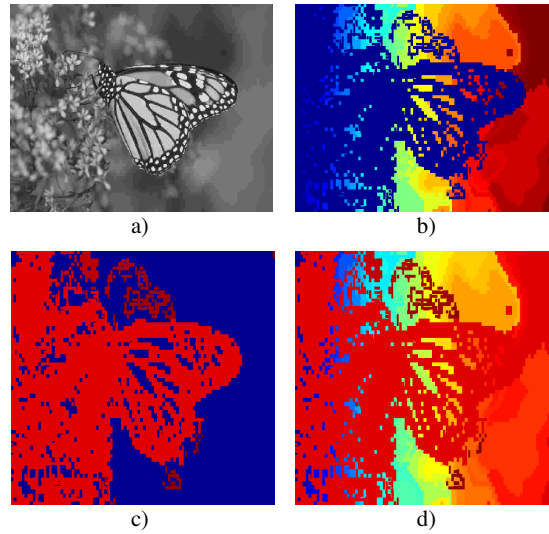


FIGURE 2.15 – a) Image compressée avec un facteur de qualité égale à 10, b) Régions de la classe C1, c) Régions de la classe C2 et d) Image segmentée.

Pour quantifier la visibilité de la transition entre deux régions (R_i et une de ses régions voisines N_{ij}), nous définissons une image contenant uniquement la région R_i avec un fond noir ainsi qu'une image contenant la région R_i et sa voisine N_{ij} . Les figures 2.16.c et 2.16.d illustrent cette représentation. Nous disposons ainsi de deux images, ce qui nous permet d'utiliser des modèles classiques du SVH. En d'autres termes, nous supposons que l'image contenant uniquement la région R_i est l'image originale. La région adjacente N_{ij} est alors considérée comme une dégradation qui vient s'ajouter au contenu initial.

La visibilité entre ce couple d'images est calculée en tenant compte des caractéristiques du SVH les plus pertinentes, à savoir la sélectivité directionnelle et fréquentielle, la sensibilité au contraste et l'effet de masquage.

Il est important de noter que la visibilité de la frontière entre deux régions adjacentes dépend non seulement de leurs informations structurelles, mais aussi du contraste local et de l'apparence globale de l'image. En d'autres termes, l'ajout d'un fond uniforme (figure 2.16.c et 2.16.d) introduit de petites variations. Nous montrerons ci-après que ces erreurs peuvent être facilement atténuées.

Pour un couple d'images donné, (région R_i et une de ses régions adjacentes

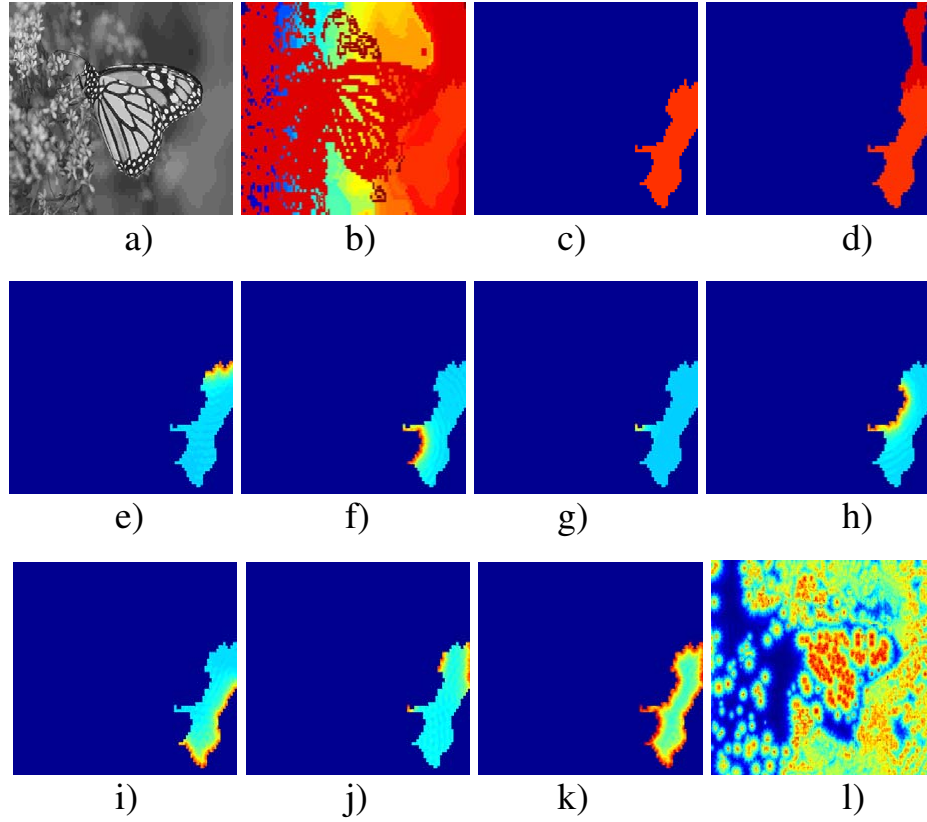


FIGURE 2.16 – a) Image compressée et sa b) version segmentée, c) Région R_i de l'image segmentée et d) avec une de ses régions adjacentes N_{ij} , e-j) Cartes de visibilité obtenues pour la R_i et chacune de ses régions voisines N_{ij} , k) Carte de visibilité obtenue pour la région R_i (toutes ses régions adjacentes) et l) Carte de visibilité globale.

N_{ij}) la carte de visibilité est obtenue comme suit :

- Filtrage CSF appliqué à chaque image.
- Décomposition multi canal (Cortex).

- Calcul du seuil d'élévation différentiel en utilisant le modèle de Daly [Dal93]^[1].
- Combinaison des cartes de visibilité de chaque sous-bande.

Les figures 2.16.e-j) illustrent les cartes de visibilité (W_i^j) obtenues pour une région donnée et chacune de ses voisines.

Ainsi, pour une région R_i donnée, m cartes de visibilité sont déduites (où m décrit le nombre de régions adjacentes). Ces cartes sont ensuite combinées comme suit :

$$W^i(x, y) = \max\{W_i^j\}, j = 1 \dots m \quad (2.7)$$

L'idée sous-jacente développée ici est de considérer pour chaque pixel, uniquement sa plus grande visibilité (principe du masquage) en fonction de son voisinage local (régions limitrophes). Comme le montre la figure 2.17.

Les figures 2.16.k et 2.17 représentent la carte de visibilité obtenue pour une région R_i donnée. On peut remarquer que les visibilités élevées sont observées à proximité des frontières entre régions. La carte de visibilité globale (V) est finalement donnée par (voir figure 2.16.l) :

$$V(x, y) = \sum_{i=1}^n W^i(x, y) \quad (2.8)$$

où n représente le nombre de régions de l'image segmentée.

Pour mieux illustrer les résultats, la méthode est appliquée sur une image synthétique (voir figure 2.18.a). Les figures 2.18.b et 2.18.c représentent la carte de visibilité des effets de bloc ainsi que son profil 1-D, respectivement. Comme on peut le constater, la carte de visibilité met clairement en évidence les frontières entre les régions adjacentes de l'image. Il est à noter aussi que l'amplitude des visibilités varie selon l'intensité des régions analysées.

[1] S. Daly. The visible differences predictor: an algorithm for the assessment of image fidelity. *Digital images and human vision*, MIT Press, 1993.

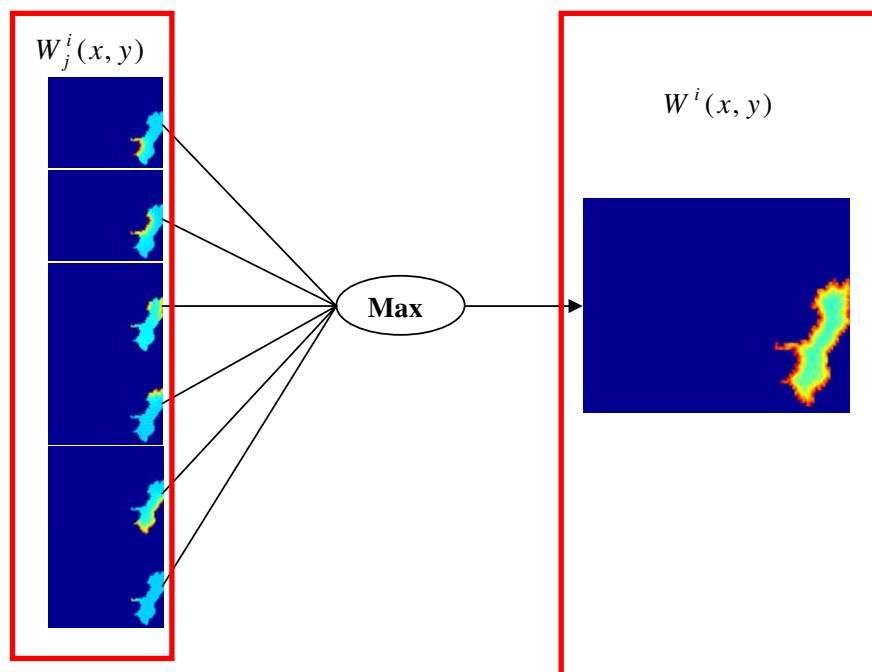


FIGURE 2.17 – Fusion des cartes de visibilité obtenues pour une région R_i donnée.

2.4.2 Réduction des effets de bloc

Plusieurs méthodes de réduction des effets de bloc ont été proposées [QDL10]^[1]. En général, les solutions ad hoc proposées visent à atténuer la transition artificielle inter bloc en utilisant des filtres de lissage. Ici, nous proposons un schéma de réduction des effets de bloc adaptatif basé sur la carte de visibilité définie précédemment. L'idée est donc d'utiliser cette carte comme entrée d'un filtre permettant ainsi d'adapter pour chaque pixel de l'image la force de filtrage.

Plusieurs filtres passe-bas peuvent être utilisés. Nous avons opté pour un filtre récursif très inspiré de celui proposé par Deriche [Der87]^[2]. L'intérêt de ce filtre est

-
- [1] Q.B. Do, A.Beghdadi and M. Luong. A New Adaptive Image Post-treatment for Deblocking and Deringing based on Total Variation method. *International Symposium on Signal Processing and Its Applications*, 2010.
 - [2] R. Deriche. Using Canny's criteria to derive a recursively implemented optimal edge de-

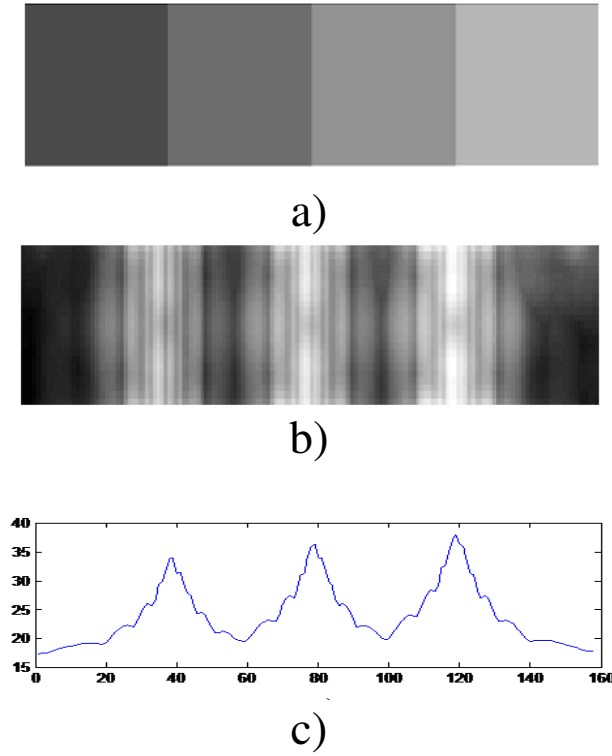


FIGURE 2.18 – a) Image synthétique, b) Carte de visibilité et son c) profil 1-D.

son caractère récursif, sa séparabilité et la possibilité d'ajuster les effets en réglant un seul paramètre.

Deux filtrages (causal et anti-causal) sont appliqués dans chaque direction :

$$y1(i) = k[x(i) + \gamma \cdot (\alpha - 1) \cdot x(i - 1)] + \gamma \cdot y1(i - 1) - \gamma^2 \cdot y1(i - 2) \quad (2.9)$$

$$y2(i) = k[\gamma \cdot (\alpha + 1) \cdot x(k + 1) + \gamma^2 \cdot x(i + 2)] + 2 \cdot \gamma \cdot y2(i + 1) - \gamma^2 \cdot y2(i + 2) \quad (2.10)$$

$$k = \frac{(1 - \gamma)^2}{1 + 2\alpha\gamma - \gamma^2} \text{ et } \gamma = \exp^{-\alpha} \quad (2.11)$$

où $y_1(i)$ et $y_2(i)$ représentent la sortie du filtre causal et anti-causal obtenues pour le pixel i , respectivement. k et γ sont deux paramètres qui dépendent de α .

Pour contrôler les variations des visibilitées des effets de bloc, nous proposons d'utiliser l'écart-type local de chaque pixel comme une mesure locale de la dégradation. La force de filtrage (α) est alors définie comme suit :

$$\alpha(x, y) = \beta \cdot \exp - \left(\frac{1}{2} \frac{\sigma[V(x, y)] - \mu}{\sigma} \right)^2 \quad (2.12)$$

où $\sigma[V(x, y)]$ est l'écart type local des visibilitées à la position (x, y) . μ et σ sont les paramètres de la gaussienne fixés ici à 0 et 2.5, respectivement. β définit l'intervalle de variation de α (fixé à 2.5).

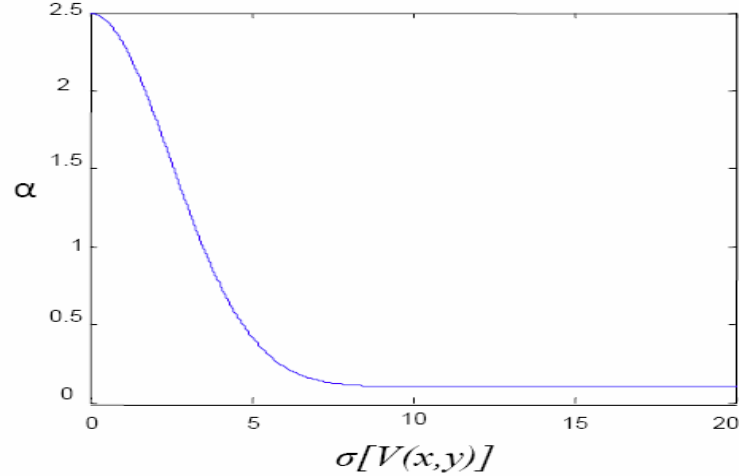


FIGURE 2.19 – Force de filtrage α en fonction de la visibilité.

Il est important de noter que l'utilisation des écarts-types permet ici de ne considérer que la distribution des coefficients de visibilité autour de la moyenne locale et ainsi d'atténuer les erreurs induites par l'ajout de fonds noirs. Une autre solution aurait été d'utiliser le contraste local ou global au lieu d'un fond noir. Toutefois, les tests expérimentaux montrent que cette solution donne de moins

bons résultats.

Pour vérifier l'efficacité de la méthode proposée, des tests expérimentaux ont été réalisés en utilisant un ensemble d'images naturelles compressées à différents taux. La procédure expérimentale est très simple et n'exige pas l'image originale ni d'information a priori concernant le type de compression ou la position des blocs. L'image corrigée est obtenue en appliquant le filtre de lissage récursif dont la force de filtrage de chaque pixel est fixée par la carte de visibilité. Cette procédure est appliquée itérativement et est arrêtée lorsque la différence entre la moyenne des visibilité à l'itération t et $t-1$ est inférieure à un certain seuil T . Les meilleurs résultats ont été obtenus pour $T=2$.

La méthode est ici comparée à celle proposée récemment par Zhai [ZZYL08]^[1] et qui représente à notre connaissance, le niveau de l'état de l'art actuel dans le domaine. Cette méthode est basée sur un filtrage suivi d'une quantification dans le domaine TCD appliqués de façons itératifs. Le critère d'arrêt est déterminé par des tests psycho-visuels. La figure 2.20 illustre les résultats obtenus sur une image test. La figure 2.20.a est l'image compressée. La figure 2.20.b est l'image obtenue après avoir appliqué la méthode de Zhai. Les figures 2.20.c et 2.20.d sont les résultats de la méthode proposée pour la 1^{er} et la 3^{ième} itérations, respectivement.

Pour une meilleure clarté, une région de l'image est aussi affichée (voir figure 2.21).

Les résultats obtenus par la méthode proposée pour une itération sont similaires à ceux obtenus par la méthode de Zhai. Pour un nombre d'itérations plus grand, de meilleurs résultats sont obtenus par la méthode proposée. De plus, nous pouvons remarquer que les régions homogènes sont plus filtrées que les régions texturées, permettant ainsi de préserver les détails de l'image et de réduire les effets de bloc.

[1] G. Zhai, W. Zhang, X. Yang, and W. Lin. Efficient Image Deblocking Based on Post filtering in Shifted Windows. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, pp. 122-126, 2008.

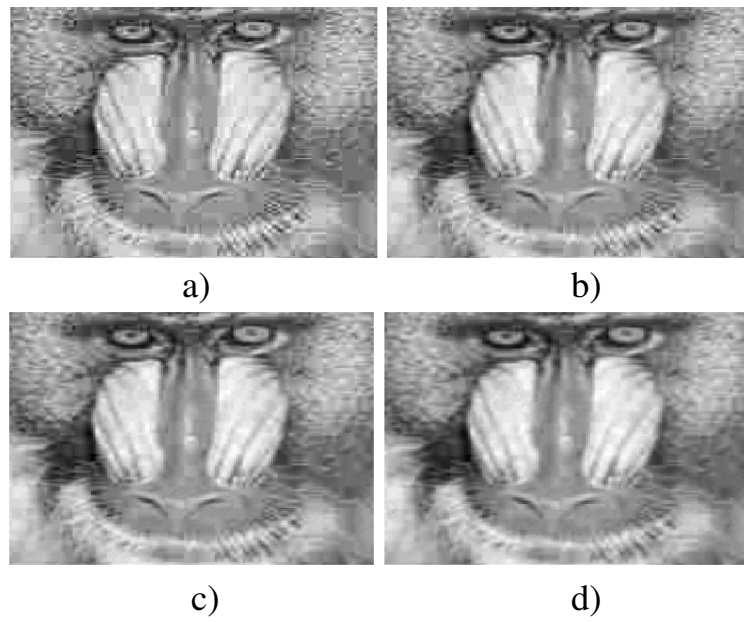


FIGURE 2.20 – a) Image compressée JPEG avec une facteur de qualité fixé à 10, b) Image filtrée par la méthode proposée par Zhai et c-d) Images filtrées par la méthode proposée pour la première et troisième itérations, respectivement.

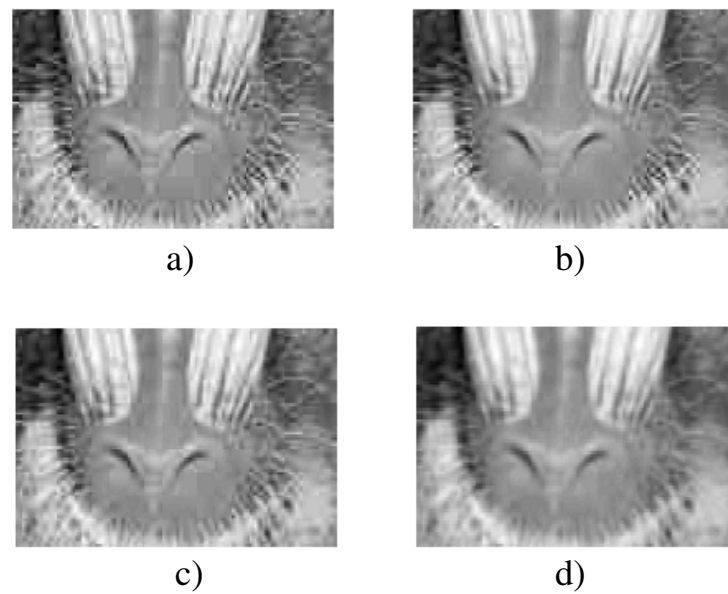


FIGURE 2.21 – Régions correspondantes aux figures 11.a, b, c et d, respectivement.

2.5 Conclusions

Ce chapitre a été entièrement consacré à l'étude des effets de bloc. Nous avons d'abord décrit la dégradation et ses origines. Nous avons ensuite présenté un bref état de l'art des méthodes d'estimation sans référence de cet artefact. La figure 2.22 résume les solutions proposées dans ce chapitre.

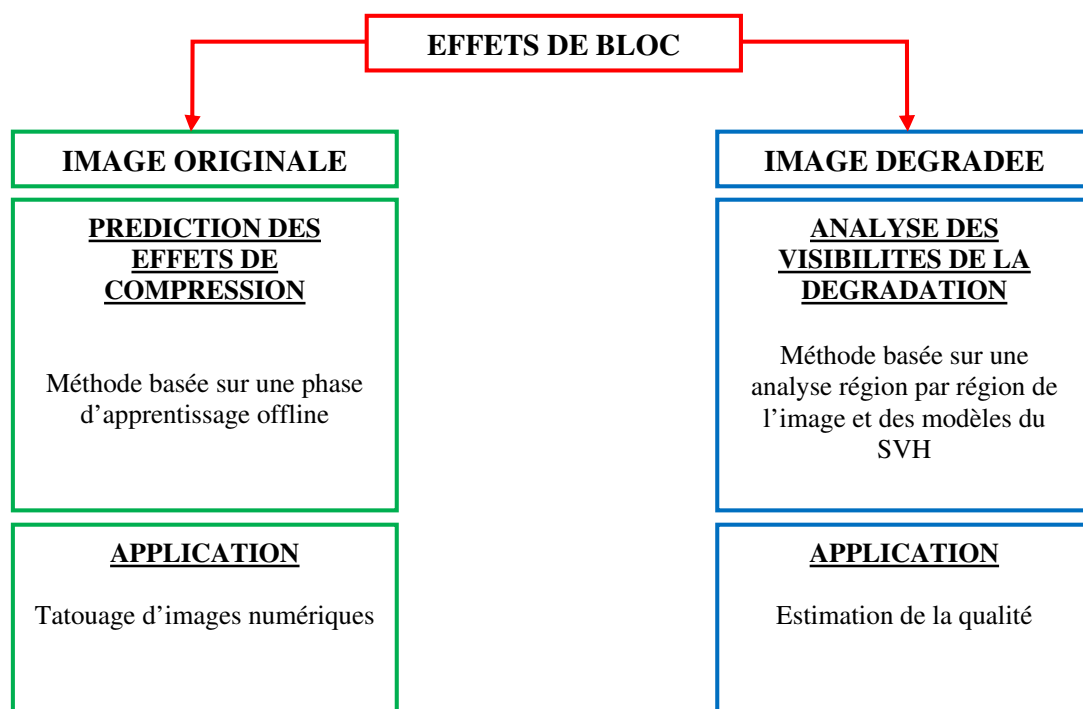


FIGURE 2.22 – Schéma récapitulatif des solutions proposées.

Prédiction des effets de blocs :

Nous avons proposé une méthode simple permettant de prédire l'apparition des effets de blocs à partir de l'image originale (non compressée). Cette méthode est basée sur une analyse locale de l'image et un apprentissage. Les résultats obtenus montrent l'efficacité de l'approche proposée. La carte de prédiction a ensuite été utilisée avec succès pour améliorer la robustesse des méthodes de tatouage face aux attaques de type JPEG. Cependant, la même

stratégie peut être appliquée aux différentes méthodes de compression par bloc.

Estimation de la visibilité des effets de bloc :

Une méthode d'estimation des effets de bloc sans référence a aussi été proposée. Cette méthode est basée sur une analyse perceptuelle par région de l'image dégradée à l'issue de laquelle une carte de visibilité est obtenue. Cette carte est ensuite exploitée pour réduire les effets de bloc au moyen d'un filtre récursif adaptatif. Les résultats obtenus confirment l'efficacité de la méthode de filtrage proposée.

La réduction de cet artefact introduit parfois un flou dans l'image. Nous proposons donc d'étudier cette dégradation dans le chapitre suivant.

Chapitre 3

Estimation du flou

Sommaire

3.1	Introduction	96
3.2	Etat de l'art	99
3.2.1	Approches spatiales	99
3.2.2	Approches fréquentielles	102
3.3	Analyse spectrale du flou	104
3.3.1	Index de qualité	106
3.3.2	Résultats expérimentaux	110
3.4	Analyse perceptuelle	111
3.4.1	Carte de visibilité perceptuelle	111
3.4.2	Index de qualité perceptuel	114
3.4.3	Résultats expérimentaux	116
3.5	Conclusions	117

3.1 Introduction

L'objectif de ce chapitre est d'étudier la dégradation de type flou qui est certainement la plus répandue et la plus gênante après les effets de bloc. Cette distorsion se manifeste essentiellement au niveau des détails et des transitions dans l'image. L'effet de lissage au niveau des contours et des textures qui en résulte affecte sensiblement la qualité de l'image par une diminution du contraste. La figure 3.1 illustre cet artefact dans le cas d'un bord net soumis à un lissage.

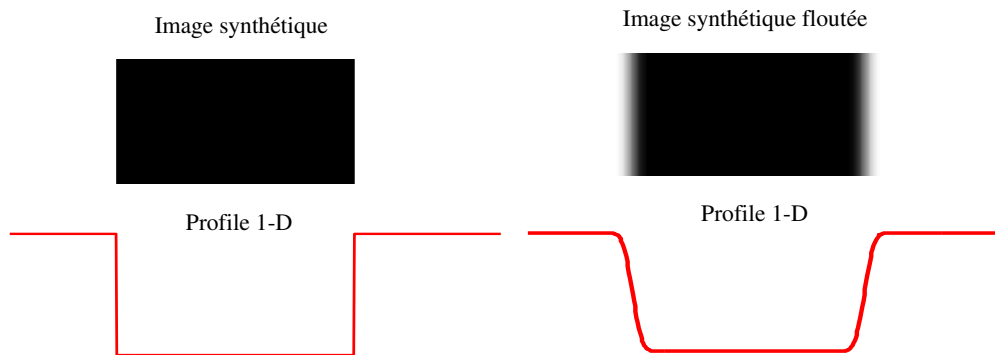


FIGURE 3.1 – Illustration d'un effet de flou sur un bord net.

Les origines de cette dégradation sont diverses et peuvent se produire à différents niveaux de la chaîne d'acquisition, de traitements et de transmission.

Flou au niveau de la capture : Le flou peut provenir d'une erreur de manipulation du matériel d'acquisition telle qu'une mauvaise focalisation ou un bougé. Le flou de défocalisation est lié à la notion de profondeur de champ qui n'est rien d'autre que l'espace d'excursion du plan de focalisation permettant d'avoir une image nette. En effet, lors d'une prise de vue, la focalisation sur un plan donné entraîne inévitablement un flou au niveau des objets se trouvant en dehors du champ de netteté. Le flou de bougé est dû essentiellement à la limitation de la vitesse d'acquisition du système de capture et de la focale. Ce flou apparaît surtout lors de l'utilisation du zoom. Ce défaut se manifeste par un effet de trainée. Cela est dû au fait que chaque élément de la scène sensibilise plusieurs pixels de la caméra. On peut

limiter cet effet indésirable en utilisant des stabilisateurs optiques. La figure 3.2 illustre un exemple de flou dû au bougé.



FIGURE 3.2 – Exemple de flou de bougé.

Mouvement : Un objet qui se déplace à grande vitesse peut produire un effet de flou sur l'image. Ce type de flou est aussi lié aux limitations du matériel de capture. On modélise ce flou par une opération de convolution d'une fonction porte orientée dans le sens du déplacement et dont l'étendue ou support n'est rien d'autre que le produit $V.T$ où V est la vitesse de déplacement et T la durée d'exposition. Un exemple de flou de mouvement est illustré par la figure 3.3.



FIGURE 3.3 – Exemple de flou de mouvement.

Perte d'informations : La perte d'information le long des canaux de transmission est aussi une cause d'apparition de flou dans l'image. La figure 3.4 présente un exemple d'image dont la perte de données est à l'origine du flou.

Compression : La compression d'image peut aussi engendrer du flou. En effet, le principe de la compression irréversible est de retenir les composantes les plus énergétiques et visibles. Les détails qui se situent en général dans les

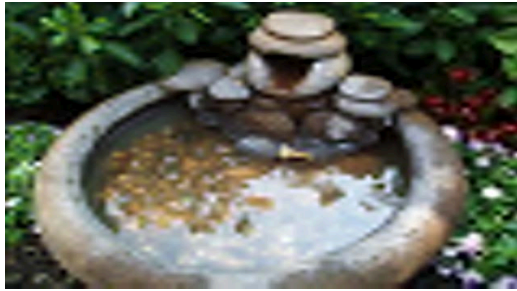


FIGURE 3.4 – Flou dû à une perte d'information.

hautes fréquences sont sacrifiées conduisant ainsi à une perte de finesse et de netteté qui se traduit par un flou dû à cet effet de filtrage passe-bas.



FIGURE 3.5 – Flou dû à la compression JPEG.

Traitements : Certains traitements visant à atténuer quelques distorsions tels que le filtrage du bruit peuvent engendrer du flou. Un exemple d'apparition de flou résultant de l'application d'un filtrage passe bas est illustré par la figure 3.6.



FIGURE 3.6 – Un flou dû à un filtrage.

Flou volontaire : Il peut aussi s'agir d'un flou volontairement ajouté à des fins artistiques, permettant ainsi de faire ressortir certaines zones de l'image au détriment d'autres. Un exemple d'image contenant un flou ajouté volontairement est présentée figure 3.7. Il est à noter qu'il semble difficile de différencier un flou artistique volontaire d'un flou indésirable. Nous considérerons donc, tout au long de cette étude, le flou comme une dégradation quelle que soit son origine.



FIGURE 3.7 – Flou artistique.

Avant de décrire les solutions proposées dans le cadre de cette thèse, nous présentons en premier lieu un bref état de l'art.

3.2 Etat de l'art

Un grand nombre de méthodes d'estimation du flou sont proposées dans la littérature. On distingue essentiellement deux approches. La première concerne les méthodes basées sur une analyse spatiale du signal. La seconde approche sont les méthodes qui agissent dans les domaines transformés tels que Fourier ou ondelettes.

3.2.1 Approches spatiales

On distingue deux types d'approches d'estimation du flou. La première repose sur un modèle qui suppose cette dégradation linéaire et invariante par translation. Le signal dégradé est lié à la version originale par une opération de convolution. Le modèle le plus utilisé est le suivant :

$$g(x, y) = f(x, y).h(x, y) + n(x, y) \quad (3.1)$$

où f est l'image originale, g sa version dégradée, h la réponse impulsionnelle qui modélise la fonction flou et n le bruit additif.

Suivant la nature de la distorsion, h peut prendre différentes formes telles qu'une gaussienne ou une fonction porte 2D. L'estimation du flou consiste alors à déterminer les paramètres de cette fonction. Il existe plusieurs méthodes avec ou sans connaissance a priori sur la dégradation (méthode classique ou aveugle) [KT97]^[1] [SAMB09]^[2].

La deuxième approche consiste à estimer le flou à partir de caractéristiques locales extraites de l'image dégradée. En général, aucune hypothèse n'est faite sur le modèle de dégradation. Les méthodes sont le plus souvent très intuitives et basées sur une connaissance a priori de l'impact du flou sur le signal. Dans ce qui suit, nous nous limiterons à cette classe de méthodes car elles permettent facilement de construire une métrique de qualité qui tient compte des structures locales de l'image, contrairement aux méthodes basées sur l'estimation de la réponse impulsionnelle qui est une caractéristique globale de la distorsion.

Dans [MDWE02]^[3], Marziliano et al. proposent d'estimer le flou en analysant la distribution des intensités des pixels à proximité des points contours. Les longueurs des contours sont estimées à partir des extrema locaux et combinées pour calculer une mesure de flou. Une version avec référence réduite a aussi été développée par les auteurs [MDWE04]^[4]. Une autre approche similaire a été proposée dans

-
- [1] D. G. Karakos and P. E. Trahanias. Generalized multichannel image-filtering structures. *IEEE Transactions on Image Processing*, Vol. 6, pp. 1038-1045, 1997.
 - [2] W. Soudene, K. Abed-Meraim and A. Beghdadi. A New Look to Multichannel Blind Image Deconvolution. *IEEE Transactions on Image Processing*, Vol. 18, pp. 1487-1500, 2009.
 - [3] P. Marziliano, F. Dufaux, S. Winkler and T. Ebrahimi. A no-reference perceptual blur metric. *International Conference on Image Processing*, 2002.
 - [4] P. Marziliano, F. Dufaux, S. Winkler and T. Ebrahimi. Perceptual blur and ringing metrics: application to JPEG2000. *Signal Processing: Image Communication*, Vol. 19, pp. 163-172, 2004.

[OLL⁺03a]^[5]. L'étalement des contours est ici analysé à partir de l'orientation des gradients.

Wang et al. proposent quant à eux d'analyser plutôt la courbure des contours de faible amplitude [WTLS08]^[6]. Dans [WX09]^[7], une mesure basée sur les contrastes et les longueurs contours est proposée.

Ferzli et al. [FK06]^[8] ont développé une métrique de qualité à partir de tests psycho-visuels dont l'objectif est d'estimer le seuil juste visible du flou (JNB : Juste Noticeable Blur) en fonction du changement de contraste. Pour ce faire, un stimulus uniforme sur fond gris est présenté à 16 observateurs. Il leur est ensuite demandé de modifier l'intensité du flou de façon à déterminer le seuil JNB. Une mesure de qualité dérivée du JNB a été développée dans [FK09]^[9].

Dans [VK08]^[10], Varadarajan et al. proposent une méthode basée aussi sur le seuil JNB. Chaque bloc (taille 64x64) de l'image est analysé et classé en bloc contour/non contour. La classification dépend ici du nombre de points contours contenus dans le bloc. La mesure de qualité globale est finalement déduite du seuil JNB et de la longueur des contours.

Une approche originale a été proposée par Crête et al. [Crê07]^[11]. Ces auteurs proposent d'estimer le flou à partir d'une analyse des variations de l'image dégradée

-
- [5] E.P. Ong, W. Lin, Z.k. Lu, X. Yang, S. Yao, F. Pan F., L. Jiang and F. Moschetti. A no reference quality metric for measuring image blur. *International Symposium on Signal Processing and Its Applications*, Vol. 1, pp. 469-472, 2003.
 - [6] X. Wang, B. Tian, C. Liang and D. Shi. Blind Image quality assessment for measuring image blur. *Congress on Image and Signal Processing*, Vol. 1, pp. 467-470, 2008.
 - [7] S. Wu and S. Xie. An efficient blind method for image quality measurement. *International Conference on Information Communications and Signal Processing*, pp. 1-5, 2009.
 - [8] R. Ferzli and L. J. Karam. A Human Visual System-Based No-Reference Objective Image Sharpness Metric. *International Conference on Image Processing*, pp. 2949-2952, 2006.
 - [9] R. Ferzli and L.J. Karam. A no reference objective image sharpness metric based on the notion of JNB. *IEEE Transactions on Image Processing*, Vol. 18, pp. 717-728, 2009.
 - [10] S. Varadarajan and L.J. Karam. An Improved Perception-Based No-Reference Objective Image Sharpness Metric Using Iterative Edge Refinement. *International Conference on Image Processing*, pp. 401-404, 2008.
 - [11] F. Crête. Estimer, mesurer et corriger les artefacts de compression pour la télévision. *Rapport de Thèse, Université Joseph Fourier*, 2007.

et de sa version filtrée. L'idée développée repose sur l'observation suivante : ” *Une image floue est moins affectée par l'ajout de flou qu'une image nette au départ*”. En d'autres termes, l'impact d'un flou ajouté à une image est différent selon que l'image contient initialement du flou ou non. Par conséquent, il suffit d'ajouter du flou à l'image et analyser son comportement face à cette nouvelle distorsion. La mesure globale est ainsi déduite de l'analyse des variations locales dans l'image filtrée par rapport à l'originale.

3.2.2 Approches fréquentielles

Contrairement aux effets de bloc, il existe une multitude de métriques d'estimation de flou sans référence exploitant les caractéristiques de la dégradation dans l'espace des fréquences spatiales.

Marichal et al. proposent une méthode sans référence basée sur une analyse des coefficients TCD [MWYJ99]^[1]. L'image est subdivisée en blocs de taille 8x8 et transformée. Le nombre d'occurrences de coefficients TCD supérieurs à un seuil donné est déterminé et comparé à un second seuil. Un poids entre 1 et 8 est alors affecté à chaque coefficient selon le nombre d'occurrences. La mesure globale est finalement obtenue en sommant l'ensemble des poids (et normalisée par la somme des poids).

Dans [CO04]^[2], Caviedes et al. proposent une mesure de flou basée sur une analyse statistique d'ordre supérieur des coefficients de la TCD [Mar74]^[3]. Le flou est ainsi estimé à partir des coefficients de Kurtosis calculés dans des fenêtres de taille 8x8 centrées aux points de contour, là où l'effet flou est le plus visible. En effet, le coefficient Kurtosis contient une information sur le degré d'aplatissement de la distribution des coefficients. En général, les images nettes ont tendance à avoir une distribution avec un coefficient Kurtosis faible. La valeur moyenne de

-
- [1] X. Marichal, M. Wei-Ying and Z. Hong Jiang. Blur determination in the compressed domain using DCT information. *ICIP, Vol. 2, pp. 386-390*, 1999.
 - [2] J. Caviedes and F. Oberti. A new sharpness metric based on local kurtosis, edge and energy information. *Signal Processing: Image Communication, Vol. 19, pp. 147-163*, 2004.
 - [3] K.V. Mardia Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya Ser, pp. 115-128*, 1974.

ces coefficients est alors utilisée comme mesure globale du flou dans l'image.

Kalalembang et al. proposent quant à eux de classer au préalable les blocs TCD en fonction de leur contenu fréquentiel [KUG09]^[4]. Ainsi, les blocs susceptibles de contenir du flou sont détectés. Un modèle de masquage est ensuite appliqué afin de déterminer si ces blocs contiennent véritablement du flou ou non.

Dans [WSH04]^[5], Wang et al. proposent une mesure d'estimation du flou basée sur l'information de phase. Une décomposition pyramidale à 3 niveaux de résolutions et 8 orientations est appliquée. L'information de phase extraite à partir des 8 sous-bandes du niveau de résolution le plus faible est utilisée pour l'estimation du flou. La même idée a été proposée dans [HWS10]^[6] où les auteurs proposent une mesure d'estimation du flou dans le domaine des ondelettes.

Tong et al. proposent aussi d'estimer la distorsion dans le domaine des ondelettes [TLZZ04]^[7]. Pour chaque niveau de décomposition, une carte de contours est d'abord déduite des coefficients hautes fréquences. Chaque carte est ensuite analysée afin de déterminer le type de contours de l'image, permettant ainsi d'estimer la dégradation.

Une étude intéressante sur la perception du flou de mouvement sur les écrans a été réalisée dans [TCB07]^[8]. Les auteurs ont mis en évidence que ce type de flou est perçu différemment selon que le flux vidéo est affiché sur un écran CRT ou LCD. L'étude est basée sur différents tests subjectifs. Les résultats de ces tests ont été utilisés pour prédire la différence de qualité entre les deux types d'écrans.

-
- [4] E. Kalalembang, K. Usman and I.P. Gunawan. DCT based local motion blur detection. *International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering*, 2009.
 - [5] Z. Wang, E.P. Simoncelli and H. Hughes. Local Phase Coherence and the Perception of Blur. *Advances Neural Information Processing Systems*, pp. 786-792, 2004.
 - [6] H. Rania, W. Zhou and S. Magdy. No reference image sharpness assessment based on local phase coherence measurement. *International Conference on Acoustics Speech and Signal Processing*, pp. 2434-2437, 2010.
 - [7] H. Tong, M. Li, Zhang and C. Zhang. Blur detection for digital images using wavelet transform. *International Conference on Multimedia and Expo, Vol. 1*, pp. 17-20, 2004.
 - [8] S. Tourancheau, P. Le Callet and D. Barba. Prediction of perceived quality differences between CRT and LCD displays based on motion blur. *International Workshop on Video Processing and Quality Metrics*, 2007.

Les lecteurs sont invités à consulter les travaux de [FCNM04]^[1] [OLL⁺03b]^[2] [CJ01]^[3] pour plus de détails.

A travers cet état de l'art, on s'aperçoit que la majorité des méthodes d'estimation de l'impact du flou sur la qualité des images sont basées sur une détection de contours. Cependant, il s'avère que cette étape préliminaire peut fortement altérer les performances de ces méthodes. En effet, une détection de contour appliquée à une image fortement floutée et contaminée par ce bruit peut s'avérer inefficace. Il serait plus judicieux de s'affranchir de cette étape de détection de contours. Nous proposons donc une approche différente inspirée de [Crê07]^[4], dont l'idée principale est d'ajouter du flou à une image donnée et d'analyser ensuite son impact.

Ainsi, deux méthodes de mesures de qualité sans référence sont ici proposées. La première méthode est basée sur une analyse spectrale. Quant à la deuxième méthode, elle est basée sur l'utilisation de modèles perceptuels et de quelques caractéristiques géométriques de la dégradation. Ces méthodes sont présentées dans ce qui suit.

3.3 Analyse spectrale du flou

L'idée principale développée ici est d'exploiter la limitation du SVH à discriminer différents niveaux de flou dans une image. La figure 3.8 met en évidence ce phénomène. La figure 3.8.a est l'image originale, ne contenant pas de flou. La figure 3.8.b est le résultat obtenu après un premier filtrage par un filtre binomial de taille 3x3. La figure 3.8.c est le résultat de deux filtrages successifs de l'image originale via le même filtre binomial 3x3. En comparant ces images, on peut re-

-
- [1] M.C.Q. Farias, M. Carli, A. Neri and S.K. Mitra A New No-Reference Method for Video Quality Assessment Using Data Hiding. *Digital Signal Processing and its Applications*, 2004.
 - [2] E.P. Ong, W. Lin, Z.k. Lu, S. Yao, X.k. Yang et L. Jinag No-reference JPEG2000. *International Conference on Multimedia and Expo, Vol. 1, pp. 545-548*, 2003.
 - [3] J. Caviedes and J. Jung. No-reference metric for a video quality control loop. *International Conference on Information Systems, Analysis and Synthesis, Vol. 13*, 2001.
 - [4] F. Crête. Estimer, mesurer et corriger les artefacts de compression pour la télévision. *Rapport de Thèse, Université Joseph Fourier*, 2007.

marquer que le flou contenu dans l'image de la figure 3.8.c n'est perceptuellement pas deux fois plus important que celui contenu dans l'image de la figure 3.8.b. En d'autres termes, l'impact d'un ajout de flou à une image déjà floue est moins important que sur une image nette au départ.

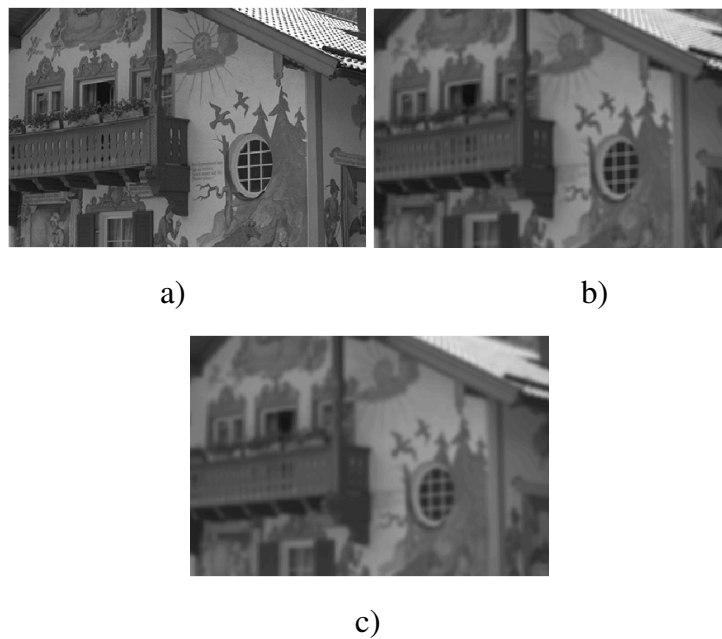


FIGURE 3.8 – a) Image originale et ses versions filtrées b) 1 fois et c) 2 fois par un filtre binomial 3x3.

Afin de mieux visualiser ce phénomène, nous considérons maintenant un signal 2D aléatoire. Nous analysons ensuite le spectre d'énergie de ce signal avant et après un filtrage passe-bas. La figure 3.9 montre le spectre d'énergie avant, après un et deux filtrages, respectivement. On peut constater que l'impact du filtrage est beaucoup plus important sur la figure 3.9.b que sur la figure 3.9.c.

Nous proposons alors d'estimer le flou selon deux méthodes qui reposent toutes les deux sur ce principe.

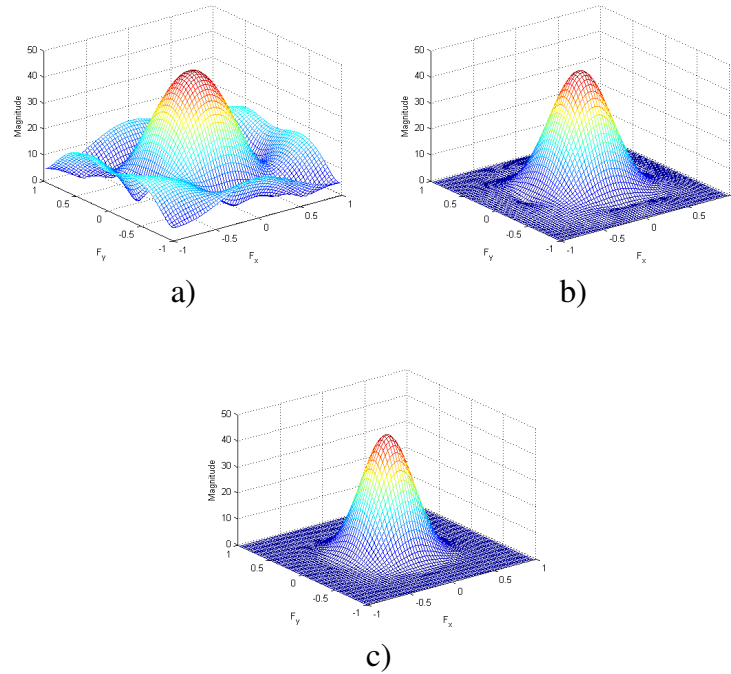


FIGURE 3.9 – Spectre d'énergie du a) signal original, b) après un filtrage, c) après deux filtrages successifs.

3.3.1 Index de qualité

Cette méthode est basée sur l'analyse de la distribution de l'énergie dans le domaine des fréquences spatiales. Le schéma synoptique est présenté figure 3.10. La première étape consiste à ajouter du flou à l'image dégradée. Une analyse de la distribution d'énergie de l'image à évaluer et de sa version lissée permet de définir un index de qualité.

L'ajout de flou est ici simulé par un filtre binomial de taille 3x3. La figure 3.11 montre un exemple d'une image et de sa version filtrée et les figures 3.11.b et 3.11.c, les versions agrandies.

La distribution spatiale du spectre d'énergie dans le plan de Fourier est analysée en utilisant la méthode décrite dans [BD00]^[1]. Cette analyse spectrale est effectuée

[1] A. Beghdadi and M. Deriche. Features extraction from fingerprints using frequency analysis. *Workshop On Signal Processing and Applications*, Vol. 18, pp. 14-15, 2000.

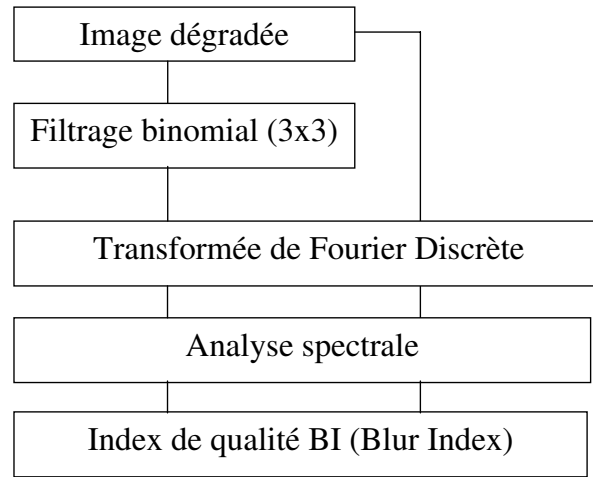


FIGURE 3.10 – Schéma synoptique.

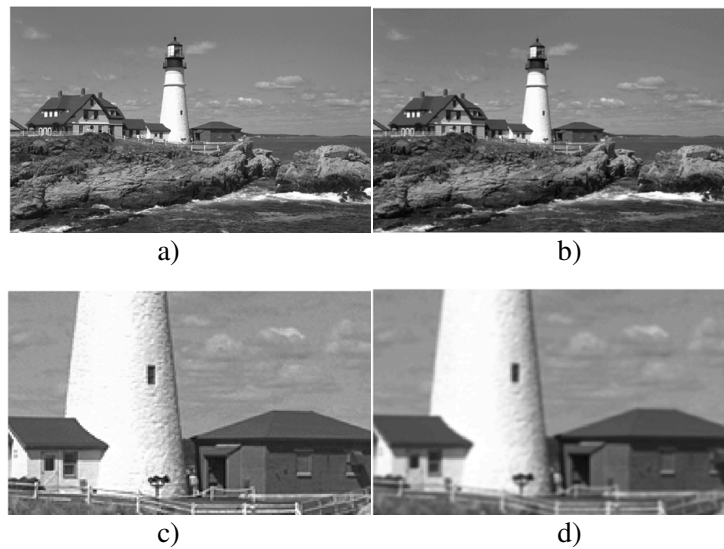


FIGURE 3.11 – a) Image test et b) sa version filtrée, zone agrandie de c) l'image test et d) de sa version filtrée.

sur l'image et sa version floue. Nous rappelons ci-après le principe de cette analyse.

Les transformées de Fourier discrète centrées de l'image ainsi que celle de sa

version filtrée sont calculées en utilisant l'expression :

$$F(u, v) = \frac{1}{X.Y} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} f(x, y) \cdot (-1)^{x+y} * \exp^{-2j\pi(\frac{ux}{X} + \frac{vy}{Y})} \quad (3.2)$$

où $F(u, v)$ est la transformée de Fourier de l'image $f(x, y)$. \mathbf{X} et \mathbf{Y} désignent le nombre de lignes et de colonnes de l'image, respectivement. \mathbf{u} et \mathbf{v} représentent les fréquences spatiales.

Le spectre d'énergie spectrale est donné par :

$$|F(u, v)| = \sqrt{R^2(u, v) + I^2(u, v)} \quad (3.3)$$

avec $\mathbf{R}(\mathbf{u}, \mathbf{v})$ et $\mathbf{I}(\mathbf{u}, \mathbf{v})$ sont les parties réelle et imaginaire de $\mathbf{F}(\mathbf{u}, \mathbf{v})$, respectivement.

A partir de la représentation en coordonnées polaire, on calcule, pour chaque fréquence ω , l'énergie totale contenue dans le disque de rayon ω (fréquence radiale) :

$$ER(\omega) = \frac{1}{K} \sum_{k=1}^K |F(\omega, \theta_k)| \quad (3.4)$$

$$ER_f(\omega) = \frac{1}{K} \sum_{k=1}^K |F_f(\omega, \theta_k)| \quad (3.5)$$

avec $ER(\omega)$ et $ER_f(\omega)$ l'énergie radiale totale du spectre de l'image test et de sa version filtrée pour la fréquence ω , respectivement. θ_k est l'angle.

La figure 3.12 présente le résultat d'analyse radiale obtenu pour une image originale et sa version floutée. Comme on pourrait s'y attendre, l'effet de flou est bien visible sur le spectre au niveau des hautes fréquences (le pic centrale correspondant à la composante continue). La variation du spectre d'énergie radiale peut donc être utilisée pour estimer l'impact de flou.

Nous proposons alors comme index du flou global la mesure donnée par 3.6.

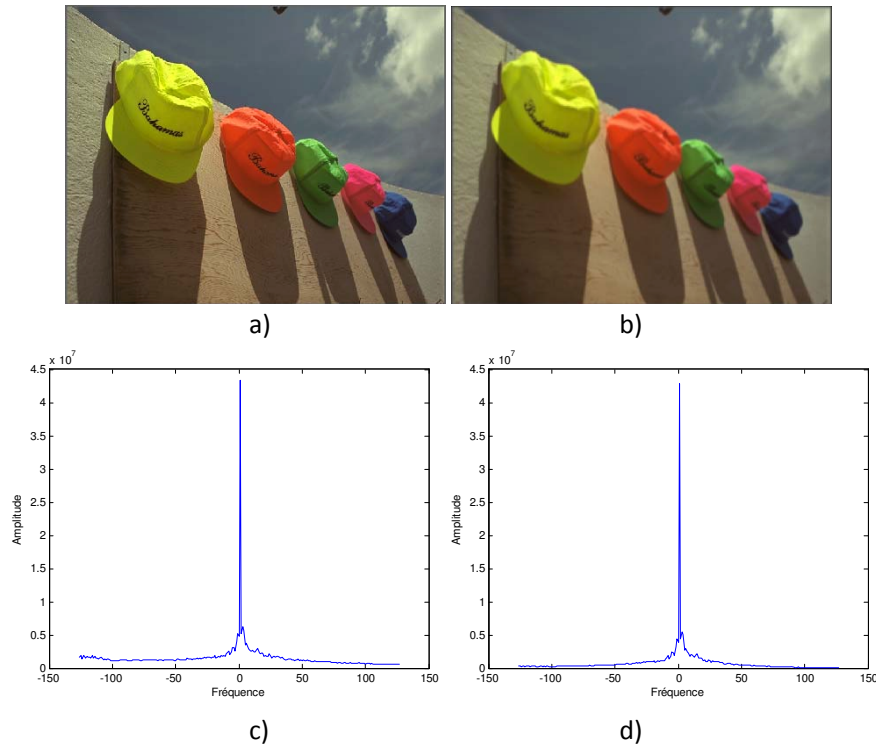


FIGURE 3.12 – a-b) Images avec différents niveaux de flou et c-d) résultats de l'analyse radiale, respectivement.

$$BI = \log\left(\frac{1}{\omega_{max}} \sum_{w=1}^{\omega_{max}} |ER(\omega) - ER_f(\omega)|\right) \quad (3.6)$$

avec ω_{max} la valeur maximale atteinte par ω .

Notons qu'il s'agit d'un flou isotrope. Dans le cas d'un bougé directionnel, il convient d'utiliser une analyse spectrale angulaire.

Méthode	LIVE	IVC
Ondelettes [TLZZ04] ^[1]	0.722	0.848
BlurM [Crê07] ^[2]	0.885	0.874
F-SSIM [WBSS04] ^[3]	0.64	0.67
Méthode proposée	0.8674	0.938

TABLE 3.1 – Corrélations obtenues.

3.3.2 Résultats expérimentaux

Nous exposons dans cette section les résultats obtenus par la méthode proposée. Elle est comparée à différentes mesures sans référence [TLZZ04]^[1] [Crê07]^[2] et à une mesure de qualité avec référence. La mesure avec référence étant ici utilisée pour estimer l'impact du flou ajouté sur l'image test. Nous choisissons pour cela, l'une des mesures les plus utilisées SSIM [WBSS04]^[3], appelée ici F-SSIM.

Le tableau 3.1 affiche les résultats obtenus en termes de corrélation avec les notes subjectives pour la base LIVE [SWCB06]^[4] et la base IVC [CA05]^[5]. On constate que les corrélations les plus élevées sont obtenues par la méthode proposée excepté pour la base LIVE, la métrique BlurM [Crê07]^[2] obtient une corrélation sensiblement plus élevée.

L'évolution de l'index de qualité BI en fonction des notes subjectives de la base LIVE est aussi présentée (voir figure 3.13). La trajectoire décrite semble correcte et non dispersée. L'allure du nuage de points montre bien que la métrique proposée est globalement en accord avec le MOS.

-
- [1] H. Tong, M. Li, Zhang and C. Zhang. Blur detection for digital images using wavelet transform. *International Conference on Multimedia and Expo, Vol. 1*, pp. 17-20, 2004.
 - [2] F. Crête. Estimer, mesurer et corriger les artefacts de compression pour la télévision. *Rapport de Thèse, Université Joseph Fourier*, 2007.
 - [3] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing, Vol. 13*, pp. 600-612, 2004.
 - [4] H.R. Sheikh, Z. Wang, L. Cormack and A.C. Bovik. LIVE Image Quality Assessment Database. <http://live.ece.utexas.edu/research/quality>, 2006.
 - [5] P. Le Callet and F. Autrusseau. Subjective quality assessment IRCCyN/IVC database. <http://www.irccyn.ec-nantes.fr/ivcdb/>, 2004.

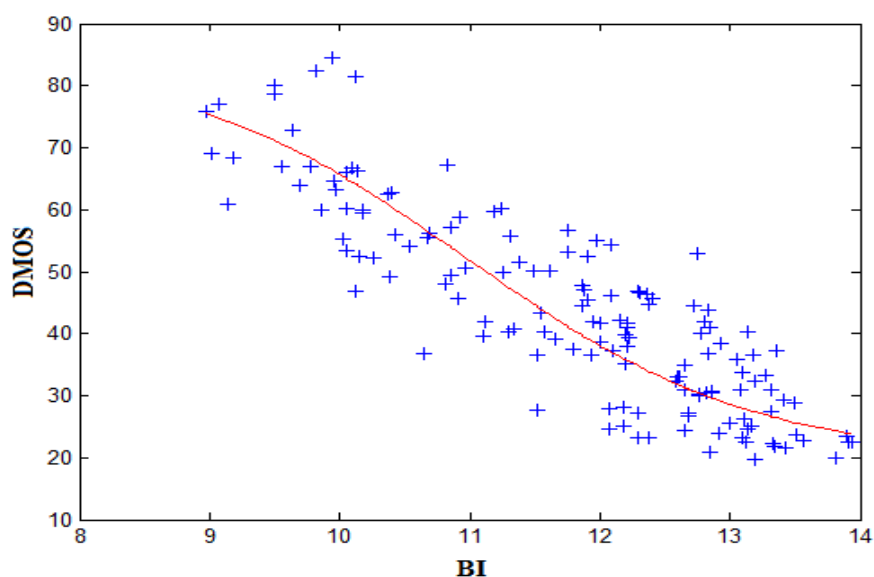


FIGURE 3.13 – Index de qualité BI en fonction du DMOS (LIVE).

3.4 Analyse perceptuelle

Nous présentons ici une méthode d'estimation du flou basée sur l'extraction d'attributs perceptuels et géométriques. Elle est basée sur le même principe que la méthode précédente en intégrant quelques caractéristiques essentielles du SVH. L'un des avantages majeur de cette méthode est de disposer, en plus d'un index de qualité, d'une carte de visibilité pouvant être ensuite utilisée, par exemple, pour réduire le flou.

3.4.1 Carte de visibilité perceptuelle

Nous proposons ici d'exploiter des modèles du SVH pour décrire, à travers une carte de visibilité perceptuelle, la quantité de flou contenue dans une image donnée. L'organigramme de la méthode proposée est présenté par la figure 3.14.

La carte de visibilité est obtenue comme suit :

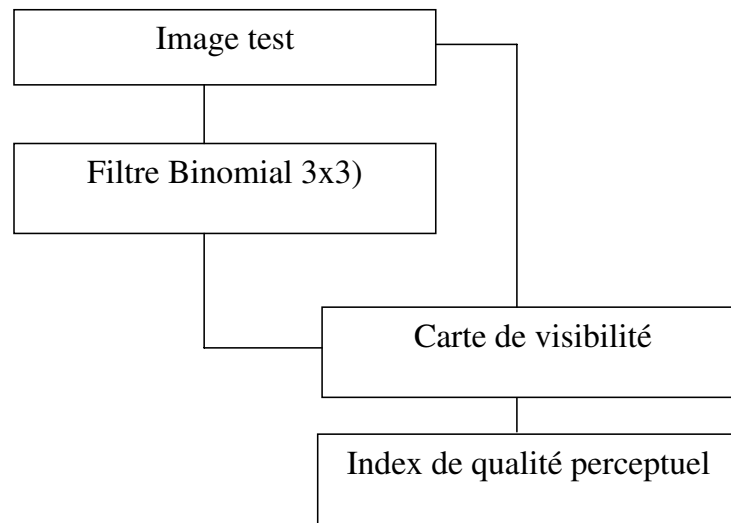


FIGURE 3.14 – Schéma synoptique.

Filtrage : L'image test est traitée par un filtre binomial de taille 3×3 . Dans [Crê07]^[1], un filtre moyenneur est utilisé. Ce filtre en plus de l'effet flou, peut affecter le contenu spatial du signal en produisant par exemple des inversions de contraste dues au caractère oscillatoire de la réponse en fréquence. Nous utilisons ici un filtre binomial de taille 3×3 .

Filtrage CSF : Pour tenir compte de la sensibilité fréquentielle du SVH, le filtre CSF est appliqué aux deux images (image test et sa version floutée). Le modèle CSF proposé par Daly est utilisé [Dal93]^[2].

Décomposition multi-canal : Une décomposition multi-canal (Cortex) est ensuite appliquée. Cette opération permet d'analyser la visibilité de la dégradation ajoutée dans chaque sous-bande de l'image.

[1] F. Crête. Estimer, mesurer et corriger les artefacts de compression pour la télévision. *Rapport de Thèse, Université Joseph Fourier*, 2007.

[2] S. Daly. The visible differences predictor: an algorithm for the assessment of image fidelity. *Digital images and human vision, MIT Press*, 1993.

Masquage : Pour mieux estimer l'impact visuel du flou dans chacune des sous-bandes, le modèle de masquage proposé par Daly est appliqué [Dal93]^[2], (une carte de visibilité par sous-bande).

Carte de visibilité : La carte de visibilité globale est alors obtenue en combinant les différentes cartes de visibilité.

Pour montrer la pertinence de la carte de visibilité face à cette distorsion, nous présentons dans la figure 3.15 une image avec différents niveaux de flou ainsi que les cartes correspondantes. On constate que les cartes de visibilité permettent bien d'estimer l'impact du flou ajouté. En effet, plus les images tests sont dégradées, moins l'impact du flou ajouté est important.

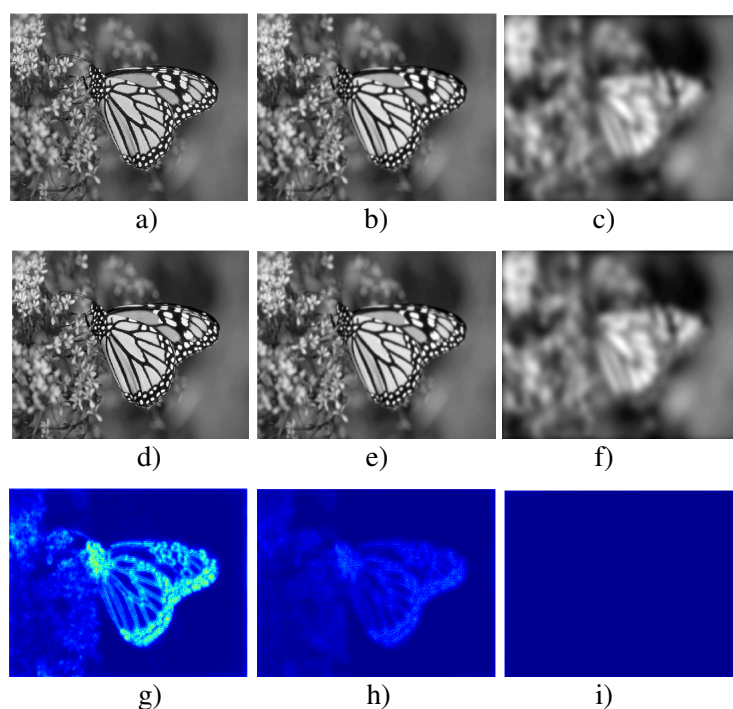


FIGURE 3.15 – a-c) Images tests et ses d-f) versions floutées ainsi que ses g-i) cartes de visibilité correspondantes.

3.4.2 Index de qualité perceptuel

Pour extraire un index de qualité à partir d'un ensemble de valeur, différentes approches peuvent être utilisées. Néanmoins, il est important de proposer une stratégie de fusion pertinente, en concordance avec le SVH. Nous proposons alors d'intégrer la notion d'étalement ou de dispersion de la dégradation qui impacte fortement la visibilité de la distorsion. En effet, de nombreuses observations confirment que les paramètres morphologiques ainsi que la distribution spatiale de la dégradation ont une grande influence sur l'évaluation de la qualité de l'image. D'autres paramètres, comme la conspécuité, la proximité, de la théorie de Gestalt [VPM96]^[1] peuvent aussi être considérés. Dans le cadre de cette étude, nous nous limitons à la distribution spatiale.

Pour extraire ou regrouper les régions de même visibilité, une segmentation par région est d'abord appliquée. Dans cette étude, nous avons opté pour une méthode de segmentation basée sur l'algorithme EM (Expectation Maximization) [Mat]^[2]. La distribution des régions de distorsions est modélisée par un mélange de gaussiennes. Cette modélisation est réalisée en deux temps : une étape d'estimation qui permet de définir les paramètres de chaque gaussienne. Une deuxième étape permet de maximiser la vraisemblance. Le nombre de classes (ou de gaussiennes) est ici déterminé automatiquement par le critère BIC (Bayesian Information Criterion). Ce critère est basé essentiellement sur une minimisation de la fonction de vraisemblance.

Un résultat de segmentation est illustré par la figure 3.16. L'image test, sa carte de visibilité et sa version segmentée y sont présentées. Il est à noter que les régions de mêmes visibilités (proches) sont bien regroupées.

Nous proposons maintenant de combiner ces informations de façon à obtenir un index de qualité unique et fiable. La notion de surface est ici prise en compte. Nous considérons ainsi que l'influence néfaste de la dégradation sur la qualité est relative à sa proportion dans l'image. En d'autres termes, plus la région dégradée

[1] V. Bruce, P. Green and M. Georgeson. Visual perception: Physiology, psychology and ecology. *LEA*, 1996.

[2] H. Zhou and H. Hu. Expectation Maximization. <http://www.mathworks.com/matlabcentral/fileexchange/10956>.

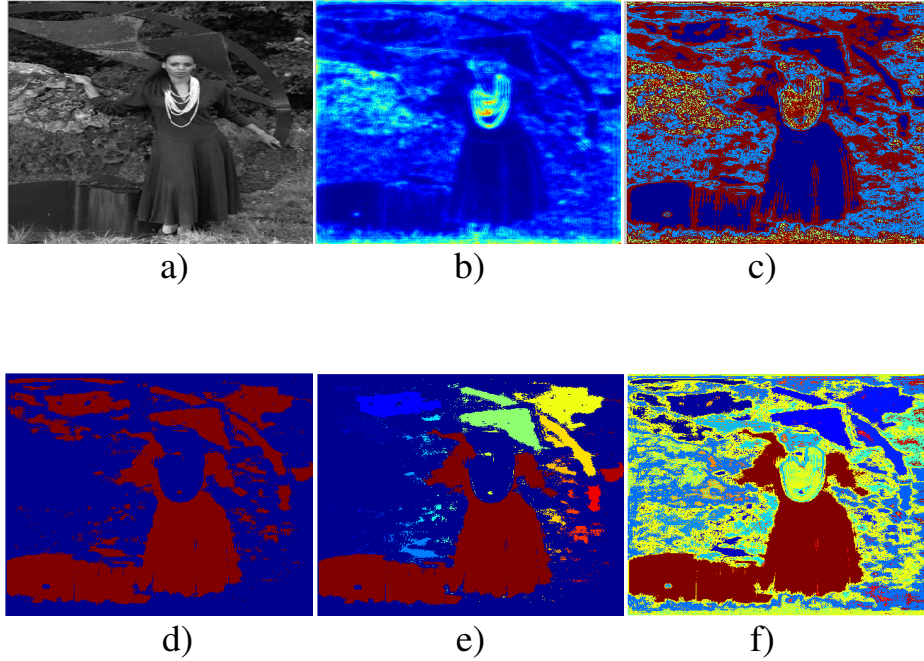


FIGURE 3.16 – a) Image test, b) carte de visibilité correspondante et sa c) version segmentée, d) Régions affectée à la classe 1 et sa e) version étiquetée, et e) version étiquetée de de l'image c) .

est grande, plus son impact est important. La figure 3.17 illustre ce phénomène (tâche grise). Cet exemple permet ainsi de mettre en évidence la nécessité de considérer la taille de la distorsion dans les schémas de fusion actuels. Ce principe a été repris et appliqué avec succès dans [SC10]^[3].

L'index de qualité est alors défini par la moyenne des visibilités de chaque région ainsi que leur taille. Nous proposons la mesure suivante :

$$PBI = \log\left(\frac{1}{R} \sum_{k=1}^R \overline{VisibMap_k} \cdot W_k\right) \text{ avec } W_k = \frac{Surface_k}{x.y} \quad (3.7)$$

où x , y et R représentent, respectivement, le nombre de lignes, colonnes et

[3] S. Chen, A. Beghdadi and A. Chetouani. A New Color Image Quality Index. *Workshop on Video Processing and Quality Metrics*, 2010.

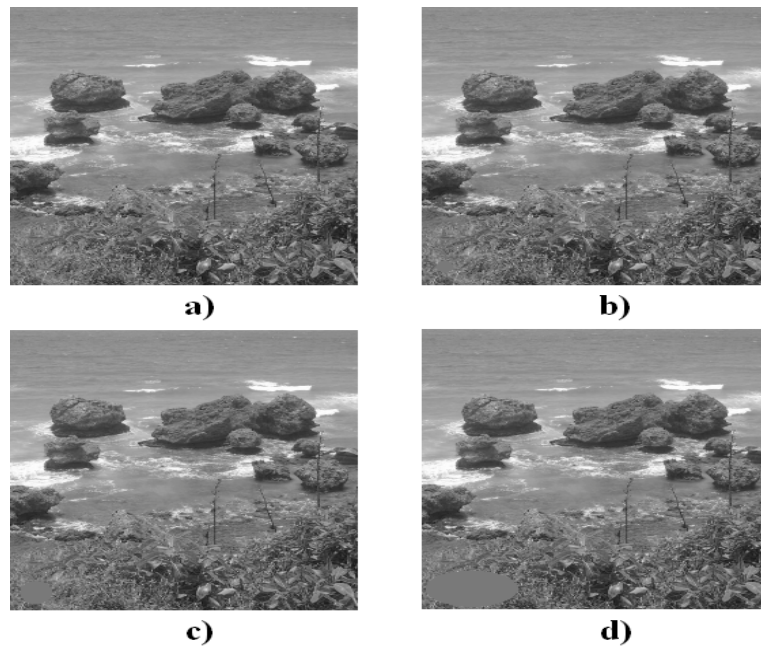


FIGURE 3.17 – a) Image originale et ses b-d) versions dégradées (tache grise).

de régions dans l'image. $\overline{VisibMap_k}$ est la visibilité moyenne de la région k . W_k correspond à la proportion (en termes de surface) de la région k dans l'image.

3.4.3 Résultats expérimentaux

La méthode proposée est comparée à différentes métriques d'estimation de flou sans référence [TLZZ04]^[1] [Crê07]^[2] et à une métrique avec référence [WBSS04]^[3].

Les résultats obtenus sont présentés dans le tableau 3.2. La meilleure corrélation est obtenue pour la méthode proposée, sauf pour la base LIVE où une corrélation sensiblement plus élevée a été obtenue par la méthode proposée par [Crê07]^[2].

-
- [1] H. Tong, M. Li, Zhang and C. Zhang. Blur detection for digital images using wavelet transform. *International Conference on Multimedia and Expo, Vol. 1*, pp. 17-20, 2004.
 - [2] F. Crête. Estimer, mesurer et corriger les artefacts de compression pour la télévision. *Rapport de Thèse, Université Joseph Fourier*, 2007.
 - [3] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing, Vol. 13*, pp. 600-612, 2004.

Méthode	LIVE	IVC
Ondelettes [TLZZ04] ^[1]	0.722	0.848
BlurM [Crê07] ^[2]	0.885	0.874
F-SSIM [WBSS04] ^[3]	0.64	0.67
Méthode proposée	0.854	0.918

TABLE 3.2 – Corrélations obtenues.

La figure 4.12 affiche la dispersion de l'index de PBI en fonction du DMOS uniquement pour la base LIVE. On observe ainsi une faible dispersion des données.

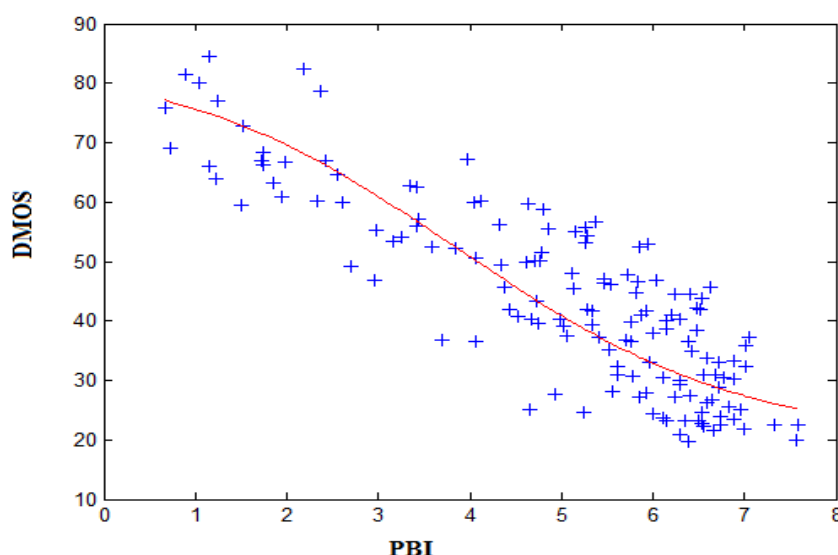


FIGURE 3.18 – Index de qualité PBI en fonction du DMOS (LIVE).

3.5 Conclusions

Dans ce chapitre, notre étude s'est focalisée sur le flou et son impact sur la qualité visuelle. Nous avons d'abord défini la dégradation ainsi que ses origines, suivi d'un bref état de l'art des méthodes d'estimation du flou sans référence. Les solutions proposées ont été ensuite exposées et décrites en détails. La figure 3.19 récapitule ces contributions.

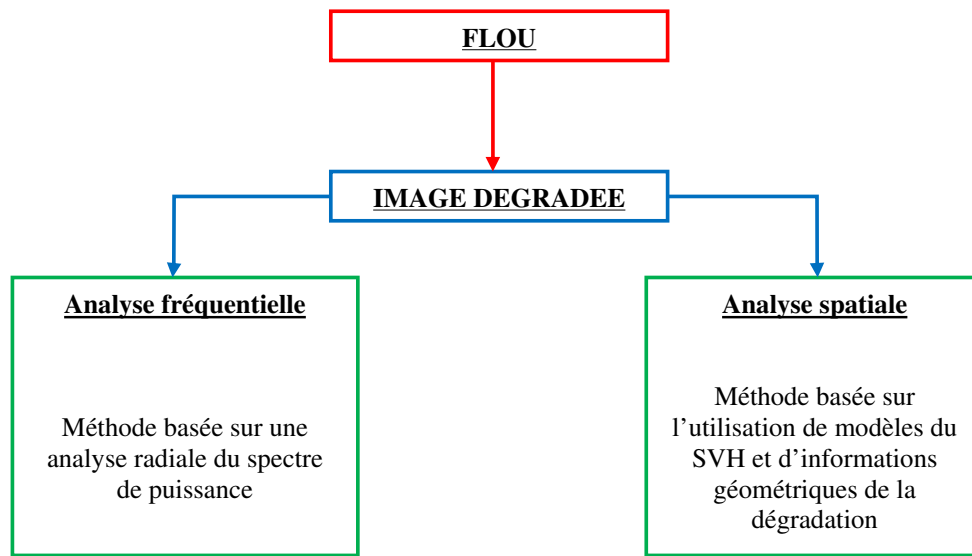


FIGURE 3.19 – Schéma récapitulatif des solutions proposées.

Analyse radiale :

Une première méthode basée sur une analyse radiale dans le domaine de Fourier a été proposée. Le principe consiste à ajouter du flou à une image et à estimer son impact. L'estimation est ici réalisée à partir d'une analyse spectrale de l'image et de sa version floue. La méthode a été testée et validée sur des bases d'images connues et utilisées par un grand nombre de chercheurs dans le domaine. Les résultats obtenus en termes de corrélation avec l'appréciation subjective ont montré la pertinence d'une telle approche. De plus, cette méthode est simple et rapide.

Estimation de la visibilité du flou :

Une deuxième méthode basée sur le même principe d'ajout de flou et quelques propriétés du SVH a été proposée. Nous avons aussi montré que la prise en compte de la distribution spatiale de la dégradation dans le calcul de la mesure du flou permet d'obtenir des résultats corrélés à l'appréciation subjective.

Chapitre 4

Estimation du ringing sans référence

Sommaire

4.1	Introduction	120
4.2	Etat de l'art	121
4.2.1	Approches spatiales	122
4.2.2	Approches dans le domaine transformé	123
4.3	Estimation du ringing sans référence	125
4.3.1	Analyse locale des gradients	125
4.3.2	Index de qualité	130
4.3.3	Résultats expérimentaux	132
4.4	Conclusion	133

4.1 Introduction

L'objet de ce chapitre est l'étude de l'effet de Gibbs 2D appelé "ringing" dans le cas du signal image. Cette dégradation est due en général à l'étape de quantification ou de décimation des coefficients hautes fréquences. Elle se manifeste sous forme d'oscillations au voisinage des régions à fort contraste et est souvent définie comme un bruit autour de ces régions. Pour mieux comprendre ce phénomène, prenons l'exemple d'un signal carré périodique. Ce dernier peut être exprimé en une somme de signaux sinusoïdaux. Lors de la reconstruction du signal, si on venait à ne considérer que certaines fréquences (sinusoïdes) nous verrions apparaître des ondulations au niveau des transitions du signal. L'amplitude des oscillations est liée à la transition (force) ainsi qu'au nombre de composantes fréquentielles retenues. La figure 5.1 illustre cet exemple où N représente le nombre de sinusoïdes considérées.

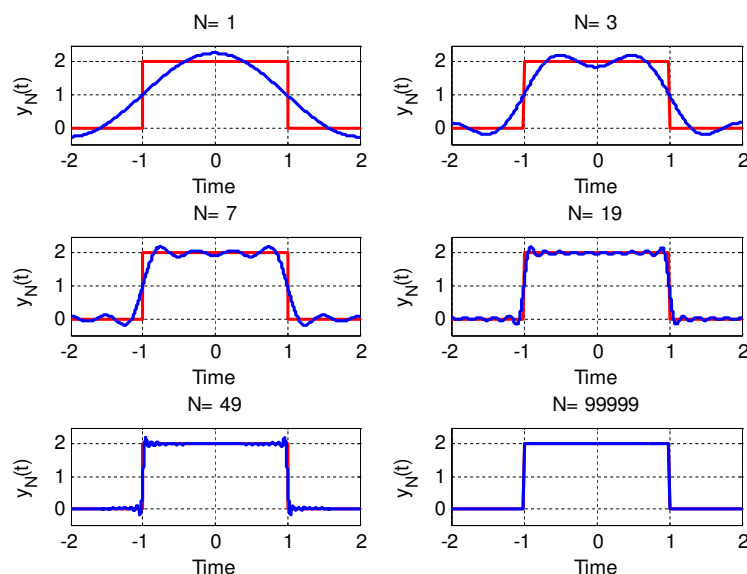


FIGURE 4.1 – Exemple d'effet de Gibbs 1D sur un signal porte.

Dans le cas des images, cet effet se traduit visuellement par l'apparition d'oscillations ou d'ombres successives le long des contours. C'est le cas notamment des

images compressées par JPEG2000 [TM01]^[1]. En effet, une fois l'image transformée dans le domaine des ondelettes, une étape de quantification est appliquée. Cette opération génère ainsi des ondulations au niveau des régions à fort contraste. Le niveau de visibilité de cette dégradation dépend donc fortement du contraste et du contenu fréquentiel. Plus le contraste est élevé, plus la distorsion incommode l'observateur. Au voisinage des zones texturées, ce phénomène est légèrement masqué mais perturbe la cohérence spatiale du signal image. A proximité d'une région homogène, la dégradation devient très gênante [ITU96]^[2] [YW98]^[3]. La figure 4.2 illustre ce phénomène dans le cas d'une image compressée par JPEG2000.



FIGURE 4.2 – Exemple de ringing dans une image.

Dans ce qui suit, nous dressons d'abord un bref état de l'art des méthodes d'estimation du ringing. Nous exposons ensuite une des limitations majeure des métriques généralement proposées et présentons notre contribution. La méthode proposée est alors décrite et évaluée à travers des résultats expérimentaux.

4.2 Etat de l'art

L'effet de ringing est de plus en plus étudié surtout depuis l'adoption du standard de compression JPEG2000, qui offre les meilleures performances en terme qualité/taux de compression.

-
- [1] D. Taubman and M. Marcellin. JPEG2000: Image compression fundamentals, standards and practice. *Kluwer Academic Publishers*, 2001.
 - [2] Principles of a reference impairment system for video. *ITU Recommendation P.930*, 1996.
 - [3] M. Yuen and H. R. Wu. A survey of hybrid MC/DPCM/DCT video coding distortions. *Signal Processing, Vol. 70*, pp. 247-278, 1998.

Certaines de ces mesures sont basées sur les caractéristiques spatiales de la dégradation. Tandis que d'autres, exploitent son contenu dans le domaine transformé. Il est à noter que nous nous limitons ici aux méthodes d'estimation sans référence. Certaines méthodes représentatives de l'état de l'art sont présentées ci-après.

4.2.1 Approches spatiales

Les méthodes spatiales exploitent généralement des mesures locales à partir de l'analyse des régions de transition telles que les contours. La méthode proposée par Oguz et al. en est un exemple [OHN98]^[1]. Les auteurs proposent de restaurer une image dégradée par filtrage morphologique. Les zones dégradées sont d'abord détectées par des opérations d'ouverture et de fermeture morphologique appliquées aux points contours. L'amplitude des oscillations est alors estimée au moyen de la variance locale. Une autre approche similaire a été proposée dans [CL05]^[2]. La mesure des oscillations est évaluée ici par le rapport entre la variance locale calculée dans un proche et lointain voisinage. Dans [BS05]^[3], les auteurs proposent quant à eux d'estimer la dégradation par le rapport de l'énergie des hautes et basses fréquences au voisinage des contours.

Dans [MDWE04]^[4], Marziliano et al. proposent de reprendre leur mesure d'estimation de flou et de l'étendre au ringing [MDWE02]^[5]. L'amplitude des oscillations ainsi que leur étalement sont d'abord mesurés et combinés pour quantifier la

-
- [1] S.H. Oguz, Y.H. Hu and T.Q. Nguyen. Image coding ringing artifact reduction using morphological post-filtering. *Workshop on Multimedia Signal Processing*, pp. 628-633, 1998.
 - [2] H. Cheng and J. Lubin. Reference free objective quality metrics for MPEG coded video. *SPIE, Human vision and electronic imaging X*, Vol. 5666, pp. 160-167, 2005.
 - [3] R. Barland and A. Saadane. Reference free quality metric for jpeg2000. *International Symposium on Signal Processing and Its Applications*, Vol. 1, pp. 351-354, 2005.
 - [4] P. Marziliano, F. Dufaux, S. Winkler and T. Ebrahimi. Perceptual blur and ringing metrics: application to JPEG2000. *Signal Processing: Image Communication*, Vol. 19, pp. 163-172, 2004.
 - [5] P. Marziliano, F. Dufaux, S. Winkler and T. Ebrahimi. A no-reference perceptual blur metric. *International Conference on Image Processing*, 2002.

dégradation. L'idée a été reprise et améliorée dans [Far04]^[6].

Une approche différente à été proposée très récemment dans [ZL10]^[7]. Zhang et al. proposent de mesurer la qualité d'une image par analyse locale des gradients. L'image est d'abord subdivisée en blocs (sans recouvrement). Une carte de l'activité locale est alors extraite des gradients de l'image et est ensuite utilisée pour le calcul d'un index de qualité. Sazzad et al. proposent une méthode similaire où l'analyse locale est limitée aux contours de l'image [SKH07]^[8].

4.2.2 Approches dans le domaine transformé

Le ringing a été aussi largement étudié dans le domaine transformé. On peut citer par exemple la méthode proposée par Sheikh et al. [SBC05]^[9] basée sur un apprentissage dans le domaine des ondelettes. Cet apprentissage permet de déterminer certains paramètres et seuils utilisés ensuite pour décrire les caractéristiques essentielles de la dégradation.

Tong et al. proposent d'analyser la distorsion par extraction de descripteurs locaux tels que les contours [TLHJZ04]^[10]. Une fenêtre est alors affectée à chaque point contour puis vectorisée et projetée dans un nouvel espace. Une projection de type ACP (Analyse en Composantes Principales) permet de discriminer entre les contours dégradés et non dégradés de l'image. Le niveau de distorsion est alors

-
- [6] M.C.Q. Farias. No-Reference and Reduced Reference Video Quality Metrics: New Contributions. *Ph.D Dissertation, Dept. of Electrical and Computer Engineering, University of California, Santa Barbara*, 2004.
 - [7] J. Zhang and T.M. Le. A new no reference quality metric for JPEG2000 images. *IEEE Transactions on Consumer Electronics*, Vol. 56, pp. 743-750, 2010.
 - [8] Z.M.P. Sazzad, Y. Kawayoke and Y. Horita. Spatial features based on no reference image quality assessment for JPEG2000. *International Conference on Image Processing*, Vol. 3, pp. 517-520, 2007.
 - [9] H. R. Sheikh, A. C. Bovik and L. K. Cormack. No-Reference Quality Assessment Using Natural Scene Statistics: JPEG2000. *IEEE Transactions on Image Processing*, Vol. 14, 2005.
 - [10] H. Tong, M. Li, Z. Hong-Jiang and C. Zhang . No reference quality assessment for JPEG2000 compressed images. *International Conference on Image Processing*, Vol. 5, pp. 3539-3542, 2004.

déterminé par l'estimation du maximum de vraisemblance.

Une mesure basée sur l'analyse des coefficients d'ondelettes a été proposée par Zhou et al. [ZXL08]^[1]. L'image est d'abord décomposée en sous-bandes (transformée en ondelettes). Les moyennes des coefficients positifs et négatifs ainsi que leurs proportions sont ensuite extraites de certaines sous-bandes spécifiques. Les coefficients d'une sous-bande donnée et sa sous-bande inférieure correspondante sont aussi comparés. Un total de 22 descripteurs sont ainsi extraits et envoyés à l'entrée d'un réseau de neurones, permettant ensuite de les combiner et d'estimer alors le niveau de la distorsion.

Il est à noter que certaines mesures de qualité existantes sont utilisées à la fois pour estimer le flou et le ringing. Comme c'est le cas notamment de la métrique proposée dans [FK09]^[2] [VK08]. En effet, les auteurs évaluent leur méthode à la fois sur des images dégradées par un flou gaussien et des images compressées par JPEG2000.

Ainsi, la majorité des mesures proposées sont basées sur une étape préliminaire de détection de contours. Cependant, l'efficacité des méthodes de détection de contours est très liée à la qualité d'image. Pour un fort niveau de distorsion, les oscillations créées peuvent être considérées comme des contours (fort contraste). L'exemple de la figure 4.3 illustre cet effet et montre que l'exploitation de l'information contour n'est pas efficace pour l'estimation du ringing.

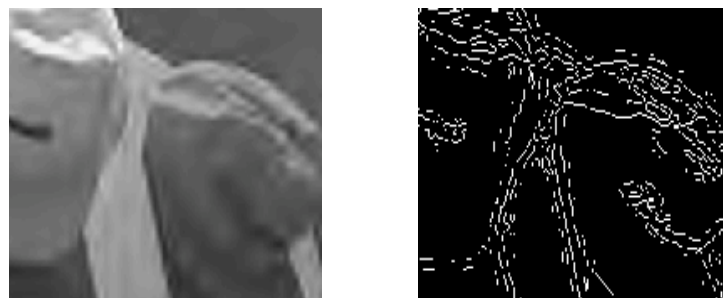


FIGURE 4.3 – Détection de contours sur une image compressée par JPEG2000.

-
- [1] J. Zhou, B. Xiao and Q. Li. A no reference image quality assessment method for JPEG2000. *World Congress on Computational Intelligence*, pp. 863-868, 2008.
 - [2] R. Ferzli and L.J. Karam. A no reference objective image sharpness metric based on the notion of JNB. *IEEE Transactions on Image Processing*, Vol. 18, pp. 717-728, 2009.

Nous proposons alors une méthode où l'on s'affranchit de cette étape de détection de contours. Dans ce qui suit, nous présentons notre contribution ainsi que les résultats obtenus.

4.3 Estimation du ringing sans référence

Dans cette étude, nous proposons d'estimer le ringing sans référence à partir d'une analyse de l'activité locale de l'image. Ce choix est lié directement à la nature de la dégradation. En effet, la dégradation se manifeste par des oscillations d'amplitude plus ou moins élevée engendrant ainsi une modification de l'activité locale du signal. De plus, cet artefact est localisé et fortement visible dans les régions à forts contraste de l'image.

L'idée développée ici est d'accentuer les variations locales des zones à fort contraste et de réduire l'activité des régions homogènes, permettant ainsi d'obtenir une carte où les contours et le ringing y sont majoritairement représentés. Le principe de la méthode proposée est illustré par la figure 4.4. L'image des gradients ainsi qu'une carte de votes des maximums locaux sont extraites de l'image dégradée. La carte de votes est ensuite projetée dans un espace paramétrique [Hou62]^[3] à partir duquel une seconde carte est obtenue. L'image des gradients est finalement pondérée par cette dernière. L'index de qualité est finalement dérivé de l'image des gradients pondérés. Ces différentes étapes sont décrites en détails dans ce qui suit.

4.3.1 Analyse locale des gradients

Le ringing crée une perturbation locale des zones à fort contraste de l'image. Pour mesurer ces variations locales, nous avons opté pour le gradient. Cependant, en plus des forts contrastes de l'image et des variations induites par le ringing, d'autres changements parasites sont détectés.

[3] P.V.C Hough. Method and means for recognizing complex patterns. *U.S. Pattern*, (3069654), 1962.

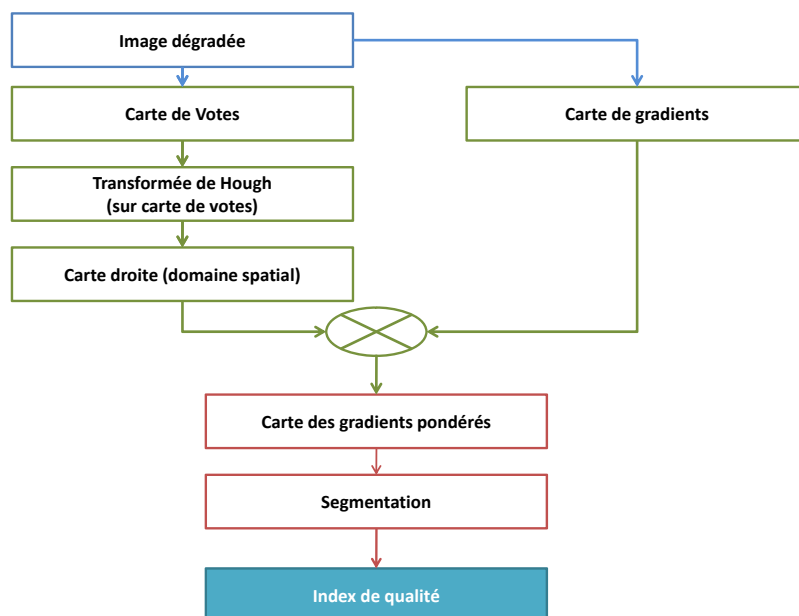


FIGURE 4.4 – Schéma synoptique.

Pour pallier ce problème, nous proposons de pondérer les gradients locaux de l'image. Cette pondération a pour objectif principal d'accentuer l'amplitude des gradients à fort contraste et de la réduire dans les zones homogènes. Cela permet alors de séparer les zones dégradées du reste de l'image.

Pour ce faire, nous proposons d'abord d'extraire une carte de votes à partir des gradients locaux de l'image. Cette carte est ici obtenue par recherche des maximums locaux dans une fenêtre glissante de taille K fixée à 11. Chaque point gradient détecté comme maximum se voit alors incrémenter de 1 dans la carte de vote (à la position correspondante). Cette opération est décrite dans la figure 4.5 et est donnée par :

$$CVote(P[\max B]) = CVote(P[\max B]) + 1 \quad (4.1)$$

Un résultat de carte de votes est présenté figure 4.6. Il est à noter que cette première étape permet de faire ressortir les contours prédominants de l'image.

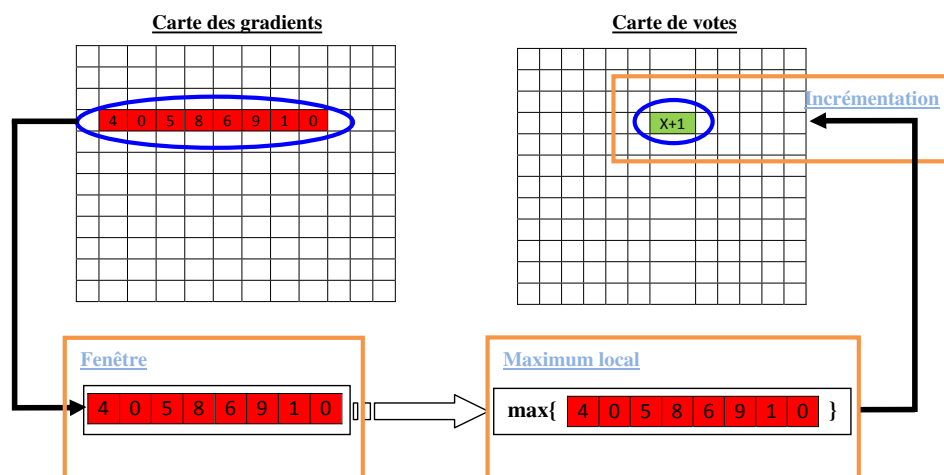


FIGURE 4.5 – Processus d'obtention de la carte de votes.

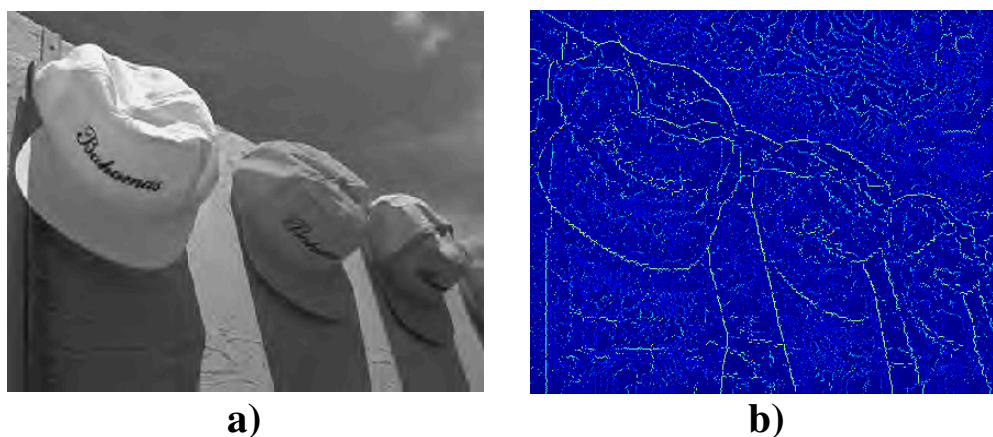


FIGURE 4.6 – a) Image dégradée et sa b) carte de votes correspondante.

Pour mieux discriminer les régions dégradées de celles non dégradées, la carte de votes est projetée à l'aide d'une transformation de Hough [Hou62]^[1] dans un espace paramétrique, à partir duquel nous déduisons une carte où les régions dégradées seront plus visibles. La transformation de Hough a été introduite en 1962 et a été très largement utilisée en traitement d'images. Elle a été initialement développée

[1] P.V.C Hough. Method and means for recognizing complex patterns. *U.S. Pattern*, (3069654), 1962.

pour détecter les droites dans une image. Le principe est le suivant :

Soit N_p un ensemble de points de coordonnées (x_i, y_i) . Chacun de ces points est transformé en une courbe sinusoïdale dans le plan paramétré (ρ, θ) .

$$\rho = x_i \cos \theta + y_i \sin \theta \quad (4.2)$$

L'ensemble des points d'une droite dans l'espace (X, Y) correspond alors à un même couple (ρ, θ) dans l'espace des paramètres (figure 4.7).

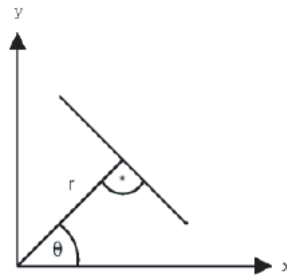


FIGURE 4.7 – Vue paramétrique d'une droite

Chaque point (x_i, y_i) vote pour une famille de droites passant par ce dernier et amène donc à une sinusoïde dans l'espace des paramètres. Idéalement, les sinusoïdes correspondant à un ensemble de points appartenant à la même droite (dans le plan (X, Y)) se croisent en un point unique dans l'espace des paramètres. La recherche de droites (dans l'image) va alors consister à retrouver dans la table représentant les couples (ρ_d, θ_d) (appelée *table des accumulations*), les points de forte accumulation (intersection des sinusoïdes). Il est à noter que dans cette étude nous avons utilisé la transformée de Hough classique. L'utilisation de la version généralisée permettrait sans doute l'amélioration des performances.

La transformée est ici appliquée sur l'ensemble des points de la carte de votes. Chaque point va alors voter dans l'espace de Hough. Le poids du vote est ici fixé par sa valeur dans la carte, permettant ainsi d'accroître la disparité entre les régions dégradées et non-dégradées. On affecte alors un poids fort aux pixels de contours ; tandis que les pixels des zones uniformes se voient attribuer un poids faible. Il est à noter que chaque point dans l'espace de Hough représente une droite dans le domaine spatial. Le processus de vote est réalisé comme suit :

$$Hough_{ij}(\rho, \theta) = Hough_{ij}(\rho, \theta) + CVote(i, j) \quad (4.3)$$

avec i et j la position du pixel dans l'image. ρ et θ la position du même pixel (i, j) dans l'espace paramétrique de Hough.

La figure 4.8 présente un exemple de résultat obtenu. Une image dégradée, la carte de votes et la table de Hough associées y sont illustrés. Les points de valeurs élevées (ou faibles) correspondent alors aux points des régions à fort contraste (ou à variation lente). La dégradation est quant à elle représentée par les valeurs intermédiaires (autour des points contours).

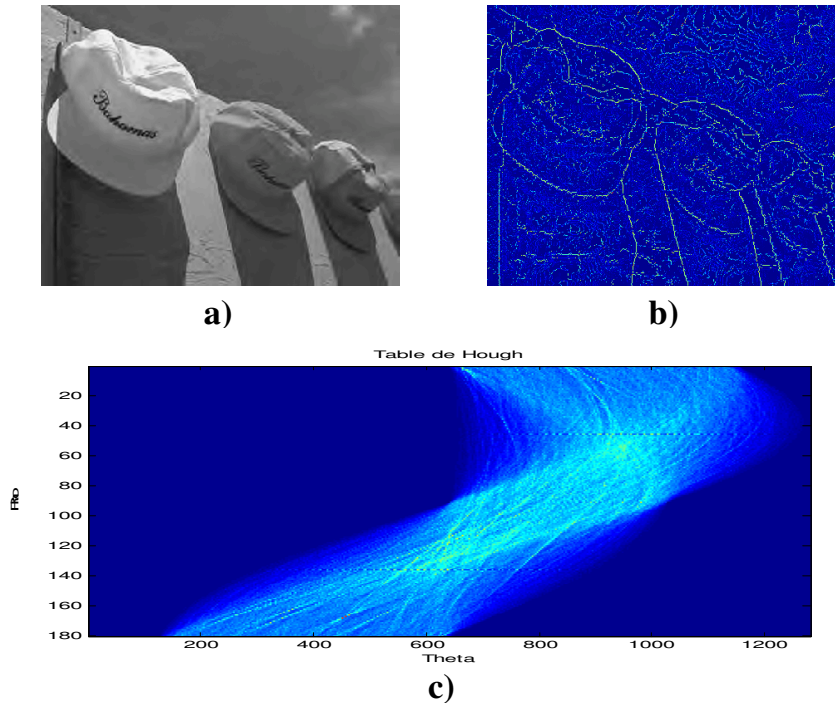


FIGURE 4.8 – a) Image dégradée, b) Carte de votes et c) sa table de Hough correspondante.

Chaque point de la table de Hough est ensuite projeté dans le domaine spatial. L'image des modules des gradients est alors pondérée par chacune de ces droites. Le poids de chaque droite est fixé par sa valeur dans la table. Cette opération permet ainsi d'atténuer (ou d'amplifier) fortement l'amplitude des gradients des régions

homogènes (ou texturées), ainsi nous séparons au mieux les régions dégradées de celles non dégradées (dans le domaine spatial).

Un résultat est présenté par la figure 4.9. L'image gradient et sa version pondérée y sont présentées. On observe que les zones correspondantes aux contours et aux régions de ringing ressortent mieux.

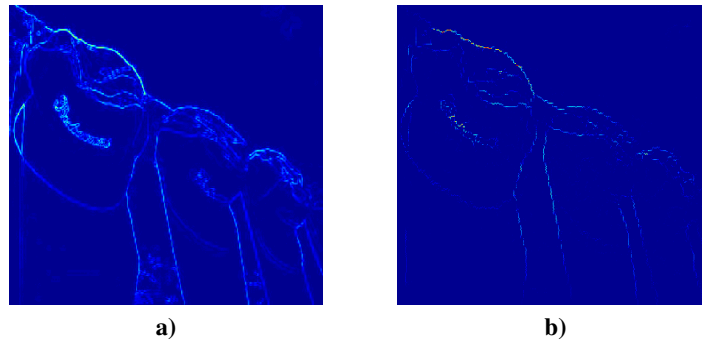


FIGURE 4.9 – a) Image des gradients et b) sa version pondérée.

La figure 4.10 résume les différentes étapes de la méthode proposée.

4.3.2 Index de qualité

A partir de la carte des gradients pondérés (CGP), un index de qualité global est déterminé en tenant compte des caractéristiques géométriques de la dégradation et de sa distribution spatiale. Ici, seule la distribution spatiale de la distorsion est prise en compte.

Pour ce faire, cette carte est d'abord segmentée, permettant ainsi de regrouper les régions de même importance et de différencier les points contours/texturés, les points ringing et les autres points (zones homogènes). Plusieurs méthodes de segmentation peuvent être utilisées. Nous utilisons ici l'approche EM (Expectation Maximization) où l'on considère la distribution des classes des points comme un mélange de gaussiennes [Mat]^[1] (voir chapitre 3). Le nombre de classes dans l'image

[1] H. Zhou and H. Hu. Expectation Maximization.
<http://www.mathworks.com/matlabcentral/fileexchange/10956>.

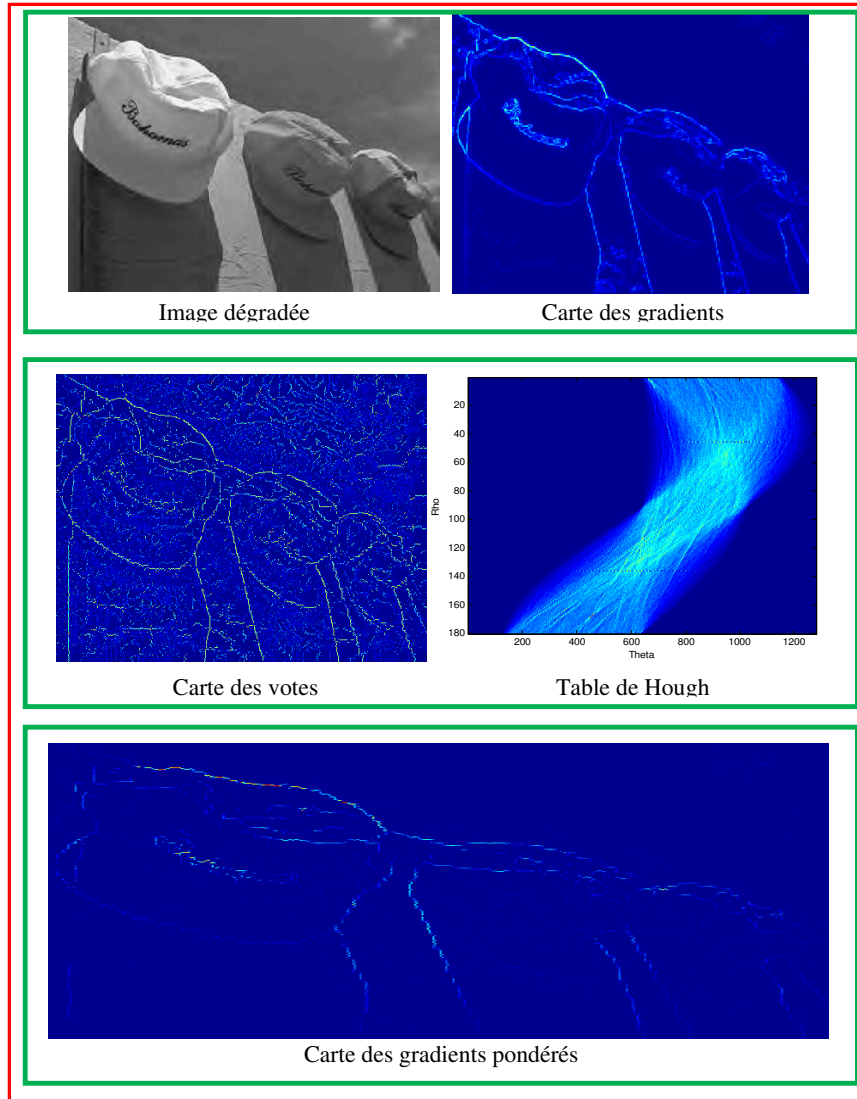


FIGURE 4.10 – Récapitulatif en images.

est déterminé par le critère BIC (Bayesian Information Criterion). La figure 4.11 présente un résultat de segmentation où l'image dégradée, la carte CGP et sa version segmentée y sont affichées.

Les régions de l'image segmentée sont alors combinées comme suit :

$$RI = \frac{1}{R} \sum_{k=1}^R \overline{GradPond_k} \cdot W_k \text{ avec } W_k = \frac{Surface_k}{x.y} \quad (4.4)$$

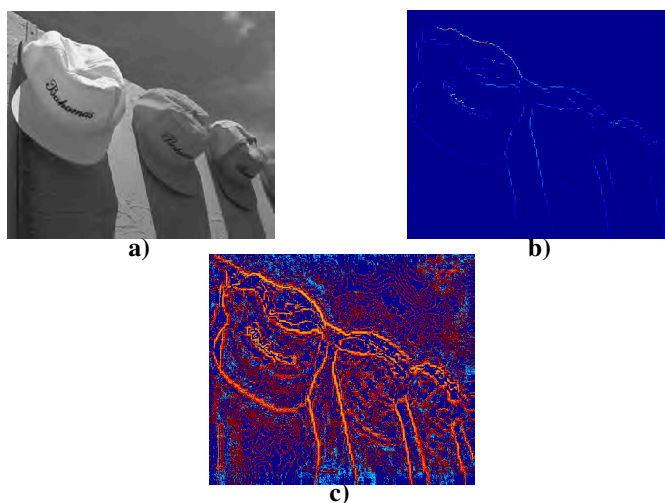


FIGURE 4.11 – a) Image dégradée, b) Carte des gradients pondérés et c) sa version segmentée.

où x , y et R représentent, respectivement, le nombre de lignes, de colonnes et de régions de l'image. $\overline{GradPond_k}$ est le gradient moyen de la région k . W_k correspond à la proportion (en termes de surface) de la région k dans l'image.

4.3.3 Résultats expérimentaux

Les performances de la méthode proposée sont comparées à 3 autres métriques de qualité sans référence : NR-QA [SBdV05]^[1], JNB [FK06]^[2] et CPBD [VK08]^[3].

Les résultats obtenus en termes de corrélation avec l'appréciation subjective sont présentées dans le tableau 4.1. Pour la base TID 2008, la méthode proposée

-
- [1] H.R. Sheikh, A.C. Bovik and G. de Veciana. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, Vol. 14, pp. 2117-2128, 2004.
 - [2] R. Ferzli and L. J. Karam. A Human Visual System-Based No-Reference Objective Image Sharpness Metric. *International Conference on Image Processing*, pp. 2949-2952, 2006.
 - [3] S. Varadarajan and L.J. Karam. An Improved Perception-Based No-Reference Objective Image Sharpness Metric Using Iterative Edge Refinement. *International Conference on Image Processing*, pp. 401-404, 2008.

Méthode	TID 2008	LIVE
JNB [FK06] ^[2]	0.81	0.71
CPBD [VK08] ^[3]	0.90	0.87
NR-QA [SBdV05] ^[1]	0.63	0.90
Méthode proposée	0.85	0.91

TABLE 4.1 – Résultats de corrélations.

obtient les meilleures performances. Cependant, pour la base LIVE, la meilleure corrélation est obtenue par la méthode CPBD.

Pour mieux observer la répartition de l'index RI par rapport aux scores subjectifs, nous affichons aussi l'index RI en fonction du DMOS (de la base LIVE) figure 4.12. On s'aperçoit que la dispersion des données est faible.

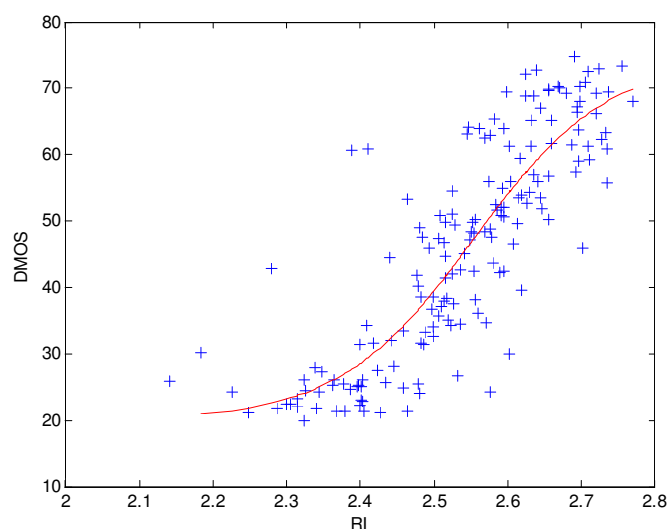


FIGURE 4.12 – Index de qualité RI en fonction du DMOS.

4.4 Conclusion

Ce chapitre a été dédié à l'analyse de l'impact du ringing sur la qualité d'image. La dégradation a d'abord été décrite ainsi que ses origines. Nous avons ensuite

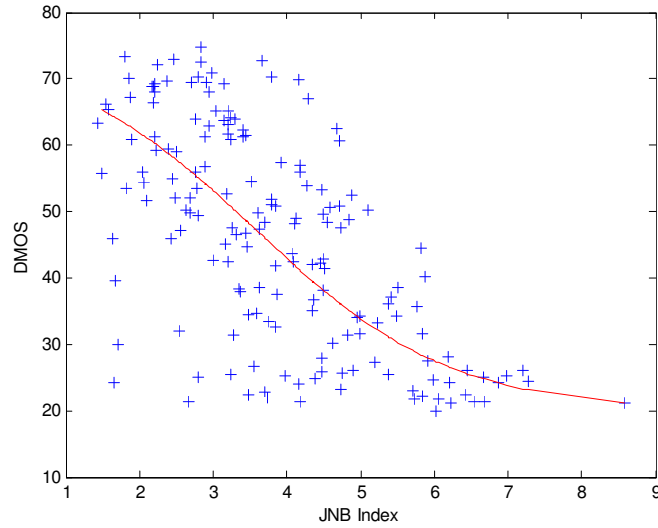


FIGURE 4.13 – Index de qualité JNB en fonction du DMOS.

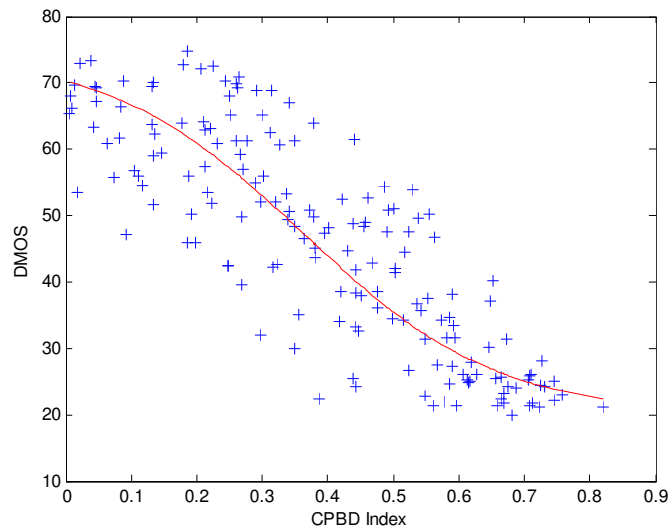


FIGURE 4.14 – Index de qualité CPBD en fonction du DMOS.

présenté un bref état de l'art des méthodes d'estimation du ringing sans référence, nous permettant alors de nous positionner et d'exposer la méthode proposée. La

figure 4.15 récapitule ces contributions.

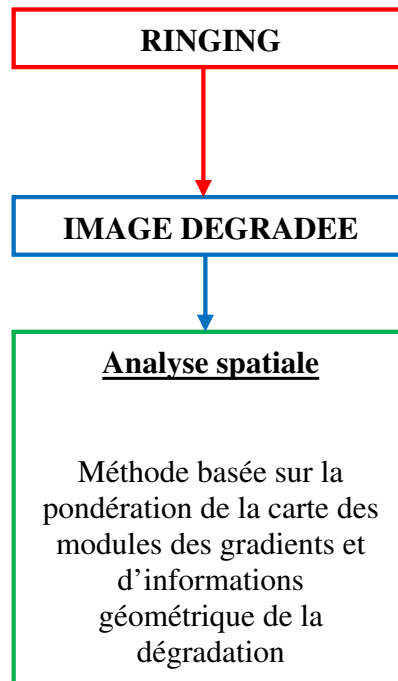


FIGURE 4.15 – Schéma récapitulatif de la solution proposée.

Nous avons proposé une méthode d'estimation de la qualité sans référence originale et efficace. Cette méthode est basée sur une analyse spatiale des gradients locaux de l'image, à l'issue de laquelle une carte des gradients pondérés est obtenue. Cette carte est ensuite utilisée pour définir un index de qualité unique décrivant la quantité de dégradation dans l'image. La méthode a été évaluée en termes de corrélation avec les notes subjectives et comparée à des métriques de qualité récentes. Les performances ainsi obtenues ont permis de valider la méthode. De plus, cette carte peut ensuite être utilisée notamment pour réduire l'impact de cet artefact.

Deuxième partie

Vers un système de mesure universel

Chapitre 5

Fusion de métriques

Sommaire

5.1	Introduction	140
5.2	Mesure de qualité avec référence	142
5.2.1	Mesures de qualité utilisées	143
5.2.2	Vers une mesure de qualité globale	149
5.2.2.1	Analyse par dégradation : <i>IQM/D</i>	150
5.2.2.2	Métrique de qualité globale (<i>G-IQM</i>)	151
5.2.3	Résultats expérimentaux	153
5.3	Mesure de qualité sans référence multi-dégradations .	155
5.3.1	NR-IQM : Mesure proposée	156
5.3.2	Résultats expérimentaux	157
5.4	Conclusions	158

5.1 Introduction

Depuis quelques années, la recherche de métriques de qualité d'image est devenue un enjeu majeur dans de nombreux domaines d'application. Une multitude de méthodes d'évaluation de qualité a ainsi vu le jour. Cependant, les mêmes hypothèses sont généralement implicitement posées.

Mesure avec référence : Universalité, utilisée quel que soit le type de dégradation.

En d'autres termes, les caractéristiques intrinsèques de la dégradation ne sont pas prises en compte.

Mesure sans référence : Non universelle, utilisée uniquement pour un type de dégradation particulier.

Il s'avère que ces hypothèses ne sont pas forcément vérifiées. L'efficacité des métriques de qualité avec référence dépend fortement du type de dégradation contenu dans l'image. En effet, les performances d'une mesure donnée sont meilleures pour certaines dégradations et moins bonnes pour d'autres.

Pour illustrer cette constatation, une expérience simple a été réalisée. Trois types de distorsions sont ici considérés : JPEG, JPEG2000 et le flou. Ces dégradations sont présentées figure 5.1. Il est à noter que cette expérience est réalisée uniquement pour les mesures avec référence.



FIGURE 5.1 – Dégradations considérées : flou (partie gauche), JPEG (partie centrale) et JPEG2000 (partie droite).

Pour chacune de ces dégradations, les métriques sont classées selon le coefficient de corrélation de Pearson (PCC). Pour cette expérience, un ensemble de me-

sure assez représentatif a été utilisé, à savoir : VIF, VIFP [SB06a]^[1], PSNR-HVS (PSNRH) , PSNRHVS-M (PSNRM) [EAP⁺06]^[2] [PSE⁺07]^[3], SSIM [WSB03]^[4], UQI [WB02a]^[5], IFC [SBdV05]^[6], WSNR [MV93]^[7], VSNR [CH07b]^[8], XYZ [KB92]^[9] et le PSNR (ces mesures sont brièvement présentées dans la section suivante).

Le classement obtenu est présenté dans le tableau 5.1. Ces résultats remettent en cause l'universalité de ce type de métriques. En effet, la mesure la plus performante pour le flou est ici VIFP, tandis que PSNRH apparaît comme la plus appropriée pour estimer la qualité des images JPEG et JPEG 2000 (pour le panel de métriques utilisées).

Les corrélations obtenues pour ces deux métriques sont présentées dans le tableau 5.2.

Ces résultats préliminaires confirment une fois de plus la nécessité de disposer d'un schéma de mesure réfléchi de façon à estimer la qualité plus efficacement.

-
- [1] H.R. Sheikh and A.C Bovik Image information and visual quality. *IEEE Transactions on Image Processing*, Vol. 15, pp. 430-444, 2006.
 - [2] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti and M. Carli. New full-reference quality metrics based on HVS. *Workshop on Video Processing and Quality Metrics*, 2006.
 - [3] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola and V. Lukin. On between-coefficient contrast masking of DCT basis functions. *Workshop on Video Processing and Quality Metrics*, 2007.
 - [4] Z. Wang, E.P. Simoncelli and A.C Bovik Multi-scale structural similarity for image quality assessment. *Asilomar Conference on Signals, Systems and Computers*, 2003.
 - [5] Z. Wang and A. Bovik A universal image quality index. *IEEE Signal Processing Letters*, Vol. 2, pp. 81-84, 2002.
 - [6] H.R. Sheikh, A.C. Bovik and G. de Veciana An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, Vol. 14, pp. 2117-2128, 2004.
 - [7] T. Mitsa and K. Varkur Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms. *International Conference on Acoustics, Speech, and Signal Processing*, pp. 301-304, 1993.
 - [8] D. Chandler and S. Hemami Subjective image database. <http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html>, 2007.
 - [9] B. Kolpatzik and C. Bouman Optimized Error Diffusion for High Quality Image Display. *Journal Electronic Imaging*, pp. 277-292, 1992.

Classement	Flou	JPEG	JPEG2000
1	VIFP	PSNRH	PSNRH
2	VIF	PSNRM	PSNRM
3	WSNR	VIF	NQM
4	VSNR	WSNR	WSNR
5	PSNRM	NQM	VIFP
6	PSNRH	VIFP	VSNR
7	SSIM	VSNR	UQI
8	UQI	SSIM	VIF
9	NQM	PSNR	PSNR
10	IFC	XYZ	SSIM
11	PSNR	UQI	IFC
12	XYZ	IFC	XYZ

TABLE 5.1 – Classement des métriques selon le coefficient de Pearson.

Dégradation	VIFP	PSNRH
Flou	0.9416	0.9120
JPEG	0.9188	0.95431
JPEG2000	0.94119	0.9568

TABLE 5.2 – Coefficient de Pearson des métriques VIF et PSNRH.

Les mesures sans référence quant à elles sont dédiées à un type de dégradation particulier. Cette donnée est généralement supposée connue, ce qui limite fortement leur utilisation dans des situations pratiques.

L'objectif de ce chapitre est de tenter d'apporter des solutions permettant de pallier ces limitations. La première partie est consacrée à l'étude des mesures de qualité avec référence. Les métriques de qualité précédemment citées y sont d'abord présentées. Le schéma de fusion proposé est ensuite défini et décrit. La deuxième partie est quant à elle dédiée à l'étude des mesures sans référence où une méthode de fusion y est présentée et discutée.

5.2 Mesure de qualité avec référence

Dans le cadre de cette étude, plusieurs métriques de qualité ont été utilisées. Chacune d'entre elles est décrite brièvement dans la section suivante.

5.2.1 Mesures de qualité utilisées

Il est à noter que la liste des métriques utilisées n'est pas exhaustive. D'autres métriques intéressantes ont aussi été proposées et peuvent être intégrées dans ce travail.

PSNR : Le rapport signal à bruit crête (PSNR) est l'une des métriques les plus connues et les plus utilisées. Elle est basée sur l'erreur quadratique moyenne et est donnée par :

$$PSNR = 10 \log\left(\frac{d^2}{MSE}\right) \quad (5.1)$$

où

$$MSE = \frac{1}{M \cdot N} \sum_{i=1}^M \sum_{j=1}^N (I_o(i, j) - I_d(i, j))^2 \quad (5.2)$$

avec $I_o(i, j)$ et $I_d(i, j)$ les intensités du pixel (i, j) de l'image originale et de sa version dégradée, respectivement. M et N représentent le nombre de lignes et de colonnes de l'image, respectivement. d est le maximum de l'intensité (255 dans le cas d'une image codée sur 8 bits).

Son succès est dû essentiellement à sa simplicité au niveau du calcul et de l'implémentation pour des applications temps réels. Cette mesure donne de bons résultats dans le cas de distorsions aléatoires et étalées dans le signal. Mais le PSNR s'avère inefficace dans le cas de dégradations structurelles localisées (non corrélée à l'appréciation subjective).

PSNR-HVS : Parmi les mesures inspirées du PSNR, la mesure récemment proposée par [EAP⁺06]^[1], appelée PSNR-HVS. La notion de sensibilité fréquentielle du SVH (CSF) est ici intégrée. L'index de qualité calculé dans le domaine TCD est donné par :

[1] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti and M. Carli. New full-reference quality metrics based on HVS. *Workshop on Video Processing and Quality Metrics*, 2006.

$$PSNR - HVS = 10 \log \left(\frac{d^2}{MSE_h} \right) \quad (5.3)$$

avec

$$MSE_h = K \sum_{i=1}^{M-7} \sum_{j=1}^{N-7} \sum_{u=1}^8 \sum_{v=1}^8 ((X[u, v]_{ij} - X[u, v]_{ij}^e) \cdot T_c[u, v])^2 \quad (5.4)$$

$$K = \frac{1}{[(M-7) \cdot (N-7) \cdot 64]} \quad (5.5)$$

avec $X[u, v]_{ij}^e$ et $X[u, v]_{ij}$ les coefficients TCD à la fréquence $[u, v]$ du bloc 8x8 de position (i, j) dans l'image originale et sa version dégradée, respectivement. M et N le nombre de lignes et colonnes de l'image, respectivement. $T_c[u, v]$ est un facteur de correction.

Une version plus récente intégrant un modèle de masquage a aussi été proposée [PSE⁺07]^[1].

Noise Quality Measure : Une mesure similaire a aussi été proposée dans [DVKG⁺00]^[2] (NQM). En plus de la CSF, le phénomène de masquage est pris en compte. L'index de qualité global est donné par

$$NQM = 10 \log \frac{\sum_x \sum_y O_s^2(x, y)}{\sum_x \sum_y (O_s(x, y) - I_s(x, y))^2} \quad (5.6)$$

où $O_s(x, y)$ et $I_s(x, y)$ est le signal original et sa version dégradée, respectivement.

Weighted Signal to Noise Rate : La qualité est ici estimée dans le domaine de Fourier [MV93]^[3]. Le spectre de puissance est pondéré par une fonction

-
- [1] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola and V. Lukin. On between-coefficient contrast masking of DCT basis functions. *Workshop on Video Processing and Quality Metrics*, 2007.
 - [2] N. Damera-Venkata, T. Kite, W. Geisler, B. Evans and A.C. Bovik Image Quality Assessment Based on a Degradation Model. *IEEE Transactions on Image Processing*, Vol. 9, pp. 636-650, 2000.
 - [3] T. Mitsa and K. Varkur Evaluation of contrast sensitivity functions for the formulation

CSF. La mesure est obtenue comme suit :

$$\text{WSNR} = 10 \log \frac{\sum_x \sum_y |X(u,v) \cdot C(u,v)|^2}{\sum_x \sum_y |(X(u,v) - Y(u,v)) \cdot C(u,v)|^2} \quad (5.7)$$

où $\mathbf{X}(u,v)$ et $\mathbf{Y}(u,v)$ sont les transformées de Fourier de l'image originale et de sa version dégradée, respectivement. $\mathbf{C}(u,v)$ est la fonction CSF utilisée.

Visual Signal to Noise Rate (VSNR) : Cette méthode basée sur l'analyse des coefficients d'ondelettes a été récemment proposée dans [CH07b]^[4]. La première étape consiste à calculer un seuil de détection dans le domaine transformé en utilisant quelques caractéristiques du SVH. L'image est considérée comme visuellement identique si le niveau de dégradation est inférieur à ce seuil. Un index de qualité est défini comme suit :

$$\text{VSNR} = 10 \log \frac{C^2(I)}{VD^2} \quad (5.8)$$

où $\mathbf{C}(I)$ est le contraste moyen de l'image originale et \mathbf{VD} est la distortion perceptuelle.

DCTune : Watson propose quant à lui d'estimer la qualité dans le domaine TCD [Wat93]^[5]. La métrique proposée est basée sur le calcul du seuil de visibilité des différences (JND). Pour chaque bloc 8x8 de l'image originale, le seuil JND est calculé à partir de l'erreur (e_{ijk}) entre l'image originale et sa version dégradée. Un modèle de masquage (m_{ijk}) est introduit dans le calcul de la carte de visibilité de la distortion. L'index de qualité est finalement donné par :

of quality measures incorporated in halftoning algorithms. *International Conference on Acoustics, Speech, and Signal Processing*, pp. 301-304, 1993.

- [4] D. Chandler and S. Hemami Subjective image database. <http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html>, 2007.
- [5] A. B. Watson DCTune: A technique for visual optimization of DCT quantization matrices for individual images. *Society for Information Display Digest of Technical Papers*, Vol. 14, pp. 946-949, 1993.

$$DCTune = \left(\sum_{i,j} p_{ij}^{\beta_f} \right)^{\frac{1}{\beta_f}} \quad (5.9)$$

avec

$$p_{ij} = \left(\sum_k \left| \frac{e_{ijk}}{m_{ijk}} \right|^{\beta_s} \right)^{\frac{1}{\beta_s}} \quad (5.10)$$

où e_{ijk} est l'erreur entre le bloc \mathbf{k} du coefficient TCD (\mathbf{i}, \mathbf{j}) de l'image originale et sa version dégradée. β_s et β_f sont deux valeurs déterminées expérimentalement.

Universal Quality Index : L'index de qualité universel (UQI) proposé par [WB02a]^[1] est basé essentiellement sur une analyse locale des moments. Il est obtenu en combinant différentes informations structurelles locales de l'image.

$$UQI = \frac{1}{W} \sum_{i=1}^W Q(i) \quad (5.11)$$

avec

$$Q(i) = \frac{4 \cdot \sigma_{xy}(i) \cdot \overline{x(i)} \cdot \overline{y(i)}}{(\sigma_x^2(i) + \sigma_y^2(i))(\overline{x(i)}^2 + \overline{y(i)}^2)} \quad (5.12)$$

avec $\sigma_x^2(i)$ et $\sigma_y^2(i)$ sont les écarts types du bloc \mathbf{i} , $\overline{x(i)}$ et $\overline{y(i)}$ sont les moyennes du bloc \mathbf{i} . W représente le nombre de blocs dans l'image. $\sigma_{xy}(i)$ est le coefficient de corrélation du bloc \mathbf{i} .

Cette métrique est très instable dans les régions homogènes (variances locales faibles dans ces régions).

Structural SIMilarity and Multi-scale SSIM : Une version améliorée de la

[1] Z. Wang and A. Bovik A universal image quality index. *IEEE Signal Processing Letters*, Vol. 2, pp. 81-84, 2002.

mesure UQI a été proposée dans [WSB03]^[2]. L'image est d'abord subdivisée en blocs de même taille. Les coefficients de corrélations de luminance, de contraste et de structure sont alors calculés pour le bloc original et son correspondant dans l'image dégradée. Un index de qualité global est déterminé par :

$$SSIM = \frac{1}{W} \sum_{i=1}^W SSIM(i) = \frac{1}{W} \sum_{i=1}^W l(i).c(i).s(i) \quad (5.13)$$

avec

$$l(i) = \frac{2.\overline{x(i)}.\overline{y(i)} + C_1}{\overline{x(i)}^2 + \overline{y(i)}^2 + C_1} \quad (5.14)$$

$$c(i) = \frac{2.\sigma_x(i).\sigma_y(i) + C_2}{\sigma_x^2(i) + \sigma_y^2(i) + C_2} \quad (5.15)$$

$$s(i) = \frac{2.\sigma_{xy}(i) + C_3}{\sigma_x(i).\sigma_y(i) + C_3} \quad (5.16)$$

$$C_1 = (L.K_1)^2, C_2 = (L.K_2)^2, C_3 = \frac{C_2}{2} \quad (5.17)$$

où \mathbf{L} est la dynamique des niveaux de gris. K_1 et K_2 sont deux constantes fixées par les auteurs à 0.01 et 0.03, respectivement.

Une version multi-échelle a aussi été proposée par les auteurs [WSB03]^[2]. Les mêmes attributs sont ici extraits à différents niveaux de décomposition permettant ainsi d'introduire l'aspect multi-échelle.

XYZ : Une métrique avec référence intégrant la CSF a été proposée dans [KB92]^[3].

Une première étape consiste à passer de l'espace RGB à l'espace XYZ. Les deux images sont ensuite filtrées par la CSF et exprimées dans l'espace CIE-LAB. Une distance ΔE est finalement calculée et utilisée comme index de qualité.

-
- [2] Z. Wang, E.P. Simoncelli and A.C Bovik Multi-scale structural similarity for image quality assessment. *Asilomar Conference on Signals, Systems and Computers*, 2003.
 - [3] B. Kolpatzik and C. Bouman Optimized Error Diffusion for High Quality Image Display. *Journal Electronic Imaging*, pp. 277-292, 1992.

$$\Delta E = \Delta L^2 + \Delta a^2 + \Delta b^2 \quad (5.18)$$

avec ΔL , Δa et Δb les erreurs entre les composantes \mathbf{L} , \mathbf{a} et \mathbf{b} de l'image originale et dégradée, respectivement.

Information Fidelity Criterion (IFC) : Basée sur l'information mutuelle, cette métrique est calculée à partir d'une modélisation de la source et de la distorsion dans le domaine des ondelettes [SBdV05]^[1]. L'index de qualité est donné par :

$$IFC = \sum_k I(C^{N_k,k}; D^{N_k,k} | s^{N_k,k}) \quad (5.19)$$

avec

$$I(C^N; D^N | s^N) = \frac{1}{2} \sum_{i=1}^N \log_2 \left(1 + \frac{g_i^2 \cdot s_i^2 \cdot \sigma_U^2}{\sigma_V^2} \right) \quad (5.20)$$

où $I(C^N; D^N | s^N)$ est l'information mutuelle conditionnelle et k désigne la sous-bande k dans le domaine ondelettes. C et D sont les modèles (mélange de gaussiennes) de l'image originale et sa version dégradée, respectivement. N représente le nombre d'éléments de chaque modèle.

Visual Information Fidelity (VIP) : Une version améliorée de la métrique IFC a été proposée dans [SB06a]^[2]. Une étape de normalisation a été intégrée. Elle est donnée par :

$$VIP = \frac{\sum_j I(\vec{C}^{N,j}; \vec{F}^{N,j} | s^{N,j})}{\sum_j I(\vec{C}^{N,j}; \vec{E}^{N,j} | s^{N,j})} \quad (5.21)$$

avec $I(\vec{C}^{N,j}; \vec{F}^{N,j} | s^{N,j})$ et $I(\vec{C}^{N,j}; \vec{E}^{N,j} | s^{N,j})$ sont les informations que l'on

-
- [1] H.R. Sheikh, A.C. Bovik and G. de Veciana An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, Vol. 14, pp. 2117-2128, 2004.
 - [2] H.R. Sheikh and A.C Bovik Image information and visual quality. *IEEE Transactions on Image Processing*, Vol. 15, pp. 430-444, 2006.

peut idéalement extraire de l'image originale et sa version dégradée à partir de la sous bande j . Une version dans le domaine spatial a aussi été proposée par les auteurs [SB06b]^[3].

5.2.2 Vers une mesure de qualité globale

Cette section est dédiée à l'étude d'un schéma de fusion efficace ayant pour objectif premier d'améliorer les performances des métriques de qualité existantes. Ces mesures doivent ainsi être utilisées de façon intelligente pour exploiter au mieux leur efficacité. Pour ce faire, nous proposons de fusionner des métriques sélectionnées au préalable. La fusion est réalisée ici en deux étapes. La première étape est liée au type d'artefact, elle consiste à combiner localement (i.e. pour chaque dégradation) certaines métriques. Une deuxième étape permet ensuite de définir la métrique de qualité globale (G-IQM). Le schéma synoptique de la méthode proposée est présenté figure 5.2.

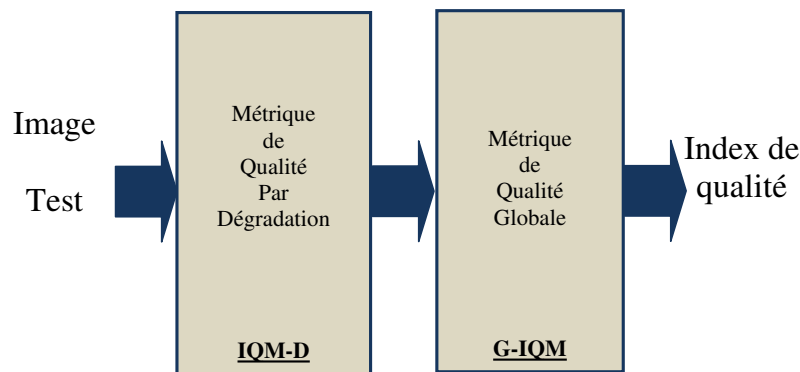


FIGURE 5.2 – Schéma synoptique.

La base d'images TID 2008 est ici utilisée. Ce choix est lié au nombre de

[3] H.R. Sheikh and A.C Bovik Image information and visual quality (in pixel domain). <http://live.ece.utexas.edu/research/quality/>, 2006.

dégradations disponibles et à la richesse des images contenues dans cette base (pour plus de détails, voir chapitre 1). Les métriques considérées dans cette étude sont listées dans le tableau 5.3.

PSNR	PNSRY	SNR
MSE	DCTune [Wat93] ^[1]	SSIM [WSB03] ^[2]
LINLAB [KB95] ^[3]	XYZ [KB92] ^[4]	PSNR-HVS [EAP ⁺ 06] ^[5]
UQI [WB02a] ^[6]	MSSIM [WSB03] ^[2]	PSNR _H VSM [PSE ⁺ 07] ^[7]
VIF [SB06a] ^[8]	VIFP [SB06a] ^[8]	NQM [DVKG ⁺ 00] ^[9]
WSNR [MV93] ^[10]	IFC [SBdV05] ^[11]	VSNR [CH07b] ^[12]

TABLE 5.3 – Métriques de qualité utilisées.

5.2.2.1 Analyse par dégradation : IQM/D

Les performances des métriques dépendent fortement du type de dégradation contenu dans l'image. Ainsi, afin de tenir compte de cette spécificité, une première étape consiste à améliorer les performances de chaque métrique en fonction de la distorsion considérée. La figure 5.3 présente la procédure de fusion (IQM/D) proposée. Pour chaque dégradation, nous sélectionnons les trois meilleures métriques en utilisant le coefficient de corrélation de Pearson comme critère de sélection. Il est à noter que le nombre de métriques sélectionnées peut être modifié. Cependant, les différents tests réalisés montrent que les performances n'augmentent pas avec le nombre de métriques.

La métrique IQM/D est alors obtenue en combinant ces métriques. La fusion est réalisée au moyen d'un réseau de neurones artificiels de type Perceptron Multi-Couche (**MLP** : Multi Layer Perceptron) décrit par la figure 5.4 dont les caractéristiques sont les suivantes :

Entrées : le nombre d'entrées du réseau de neurones est égal 3 (i.e. nombre de métriques sélectionnées).

Couche(s) cachée(s) : Pour éviter d'augmenter la complexité du système, le nombre de couches cachées est ici fixé à 1 (Le nombre de neurones diffère selon le type de dégradation).

Sortie : 1 neurone dont la cible est le MOS.

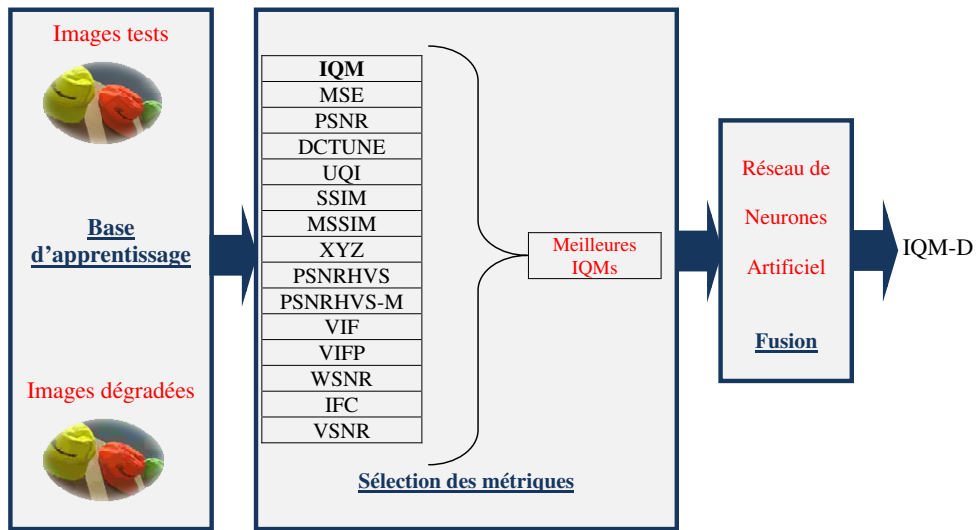


FIGURE 5.3 – IQM/D : schéma fonctionnel.

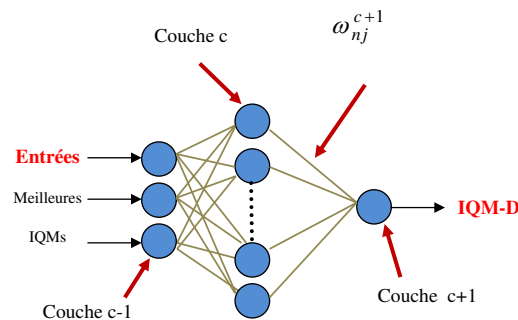


FIGURE 5.4 – IQM/D : modèle neuronale.

Fonction d'activation : Sigmoidé.

Apprentissage : Algorithme de rétro-propagation.

5.2.2.2 Métrique de qualité globale (*G-IQM*)

Un second réseau de neurones est ensuite utilisé afin d'obtenir une mesure unique quel que soit le type de dégradation. Ces entrées correspondent aux sorties des IQM-Ds (i.e. 17 entrées). Le modèle est présenté par la figure 5.5. Ses caractéristiques sont les suivantes :

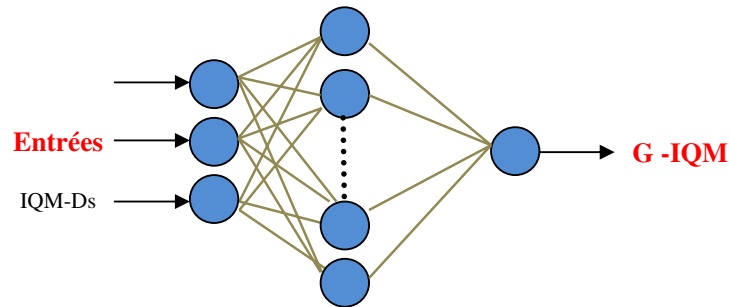


FIGURE 5.5 – G-IQM : modèle neuronal utilisé.

Entrées : 17 (i.e. sorties des 17 IQM-Ds).

Couche cachée : 1 (le nombre de neurones est égal à 10).

Sortie : 1 (La cible étant le MOS).

Fonction d'activation : Sigmoidé.

Apprentissage : Algorithme de rétro-propagation.

Ainsi, le schéma de fusion globale est illustré par la figure 5.6.

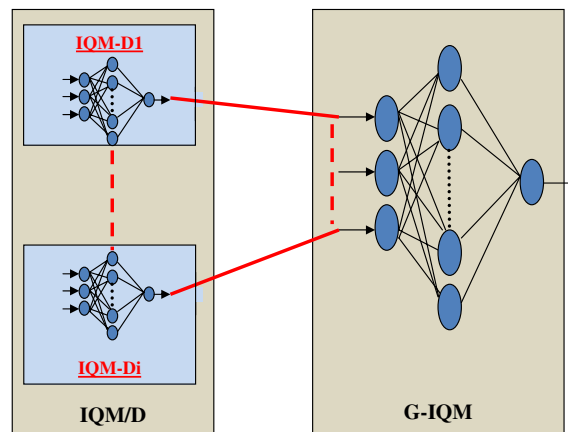


FIGURE 5.6 – Schéma de fusion globale

5.2.3 Résultats expérimentaux

Plus de 400 images ont été utilisées pour la validation du schéma de fusion proposé où toutes les dégradations y sont également représentées. Ces images sont évidemment différentes de celles utilisées lors de l'apprentissage. Le tableau 5.4 présente, pour chaque distorsion, les corrélations de Pearson obtenues par la meilleure métrique de qualité ainsi que celles obtenues pour chaque IQM/D.

Dégradation	IQM	IQM/D
1	0.57 (PSNR)	0.93
2	0.63 (PSNR)	0.92
3	0.53 (PSNRHVS)	0.95
4	0.51 (VIF)	0.81
5	0.57 (PSNRHVS)	0.98
6	0.51 (WSNR)	0.90
7	0.53 (PSNR)	0.91
8	0.48 (PSNRHVS)	0.95
9	0.61 (PSNRHVSM)	0.97
10	0.71 (PSNRHVS)	0.99
11	0.57 (WSNR)	0.97
12	0.37 (VIF)	0.92
13	0.51 (PSNRHVS)	0.92
14	0.52 (VIFP)	0.83
15	0.55 (SSIM)	0.88
16	0.54 (DCTune)	0.75
17	0.65 (VIF)	0.96

TABLE 5.4 – Corrélations de Pearson obtenues pour chaque dégradation.

Les performances obtenues par la méthode proposée sont nettement supérieures à celles des métriques existantes. Ainsi, cette première étape permet de mieux estimer la qualité des images pour chacune des dégradations considérées.

Le tableau 5.5 présente les corrélations de Pearson obtenues pour l'ensemble des métriques considérées et la mesure proposée **G-IQM**. Pour les mesures classiques, les meilleures performances ont été obtenues par la métrique MSSIM où une corrélation de 0.81 a été obtenue, tandis qu'une corrélation de 0.90 a été obtenue pour G-IQM, soit une augmentation de près de 10%.

Les figures 5.7 et 5.8 présentent la distribution de l'index G-IQM et de la

Métrique	Corrélation de Pearson	Métrique	Corrélation de Pearson
PSNR	0.50	PSNRY	0.53
SNR	0.51	MSE	0.31
DCTUNE	0.27	UQI	0.69
SSIM	0.64	MSSIM	0.81
LINLA	0.25	XYZ	0.52
PSNRH	0.58	PSNRM	0.55
VIF	0.79	VIFP	0.65
NQM	0.61	WSNR	0.46
IFC	0.21	VSNR	0.27
G-IQM	0.90		

TABLE 5.5 – Corrélations de Pearson obtenues pour ***G-IQM***

métrique MSSIM en fonction des notes subjectives (toutes dégradations confondues). On s'aperçoit que la dispersion des données de la mesure proposée est moins importante que celles obtenues par MSSIM.

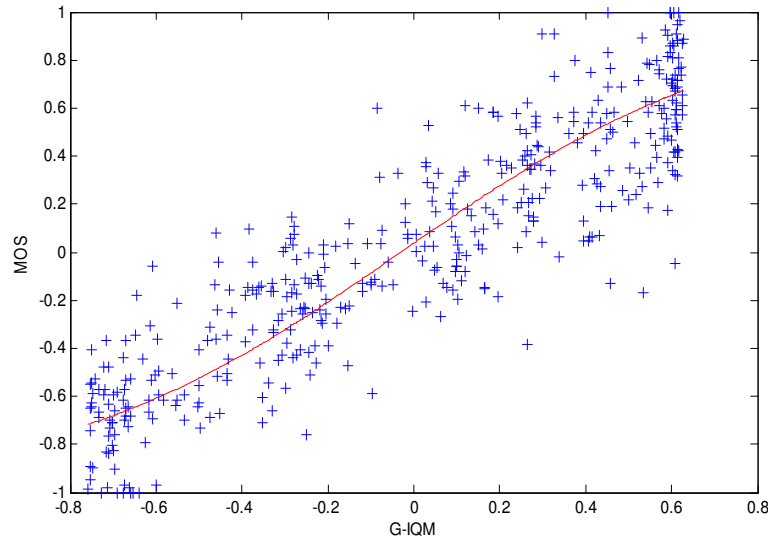


FIGURE 5.7 – Index de qualité G-IQM en fonction du MOS.

Pour mieux visualiser les résultats obtenus, la figure 5.9 présente les index de qualité pour un échantillon d'images (en rouge : MOS, bleu : G-IQM). Les valeurs de sorties (G-IQM) et le MOS varient de façon similaire (quasi).

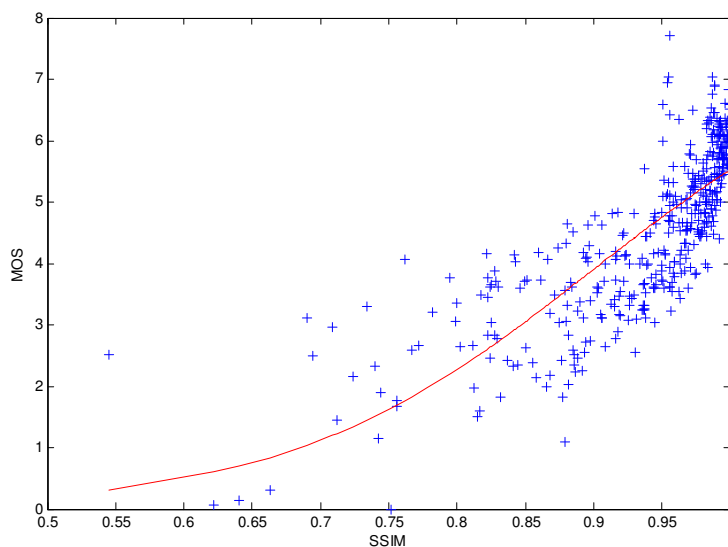


FIGURE 5.8 – Index de qualité SSIM en fonction du MOS.

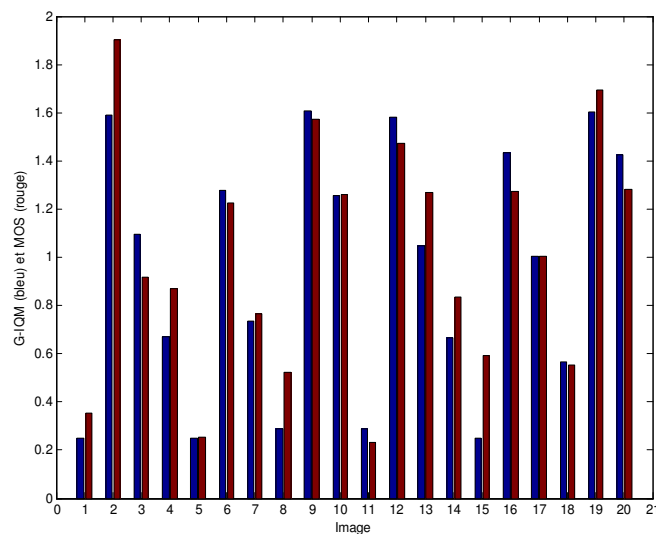


FIGURE 5.9 – Index de qualité pour différentes images : MOS (rouge) et G-IQM (bleu).

5.3 Mesure de qualité sans référence multi-dégradations

Dans un cadre applicatif, les mesures sans référence sont les plus attractives. Elles permettent de mesurer la qualité sans aucune information a priori sur l'image

originale. Cependant, ce type de mesures est orienté vers une dégradation qui est supposée connue. Leur utilisation est alors fortement limitée.

5.3.1 NR-IQM : Mesure proposée

Dans cette section, nous proposons de pallier ce problème en proposant une métrique sans référence multi-dégradations basée sur la fusion de plusieurs mesures. Il est à noter que notre étude se limite ici aux dégradations les plus répandues, à savoir les effets de blocs, le flou et le ringing. Le schéma synoptique est illustré par la figure 5.10.

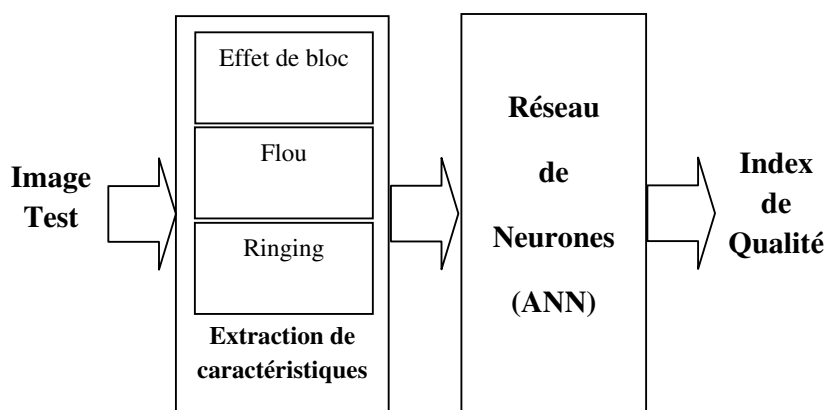


FIGURE 5.10 – Schéma synoptique.

Les métriques suivantes ont été utilisées pour la conception de la mesure.

Effets de bloc : Pour la mesure des effets de bloc, nous avons opté pour celle proposée par Wang et al. [WBE00]^[1]. Cette métrique est basée sur une analyse verticale et horizontale des frontières des blocs dans le domaine spatial.

Flou : Pour le flou, nous considérons la mesure décrite au chapitre 2 qui est basée sur l'analyse radiale du spectre de puissance de l'image [CBD09]^[2].

-
- [1] Z. Wang, A.C. Bovik and B.L. Evans. Blind measurement of blocking artefacts in images. *International Conference on Image Processing*, 2000.
 - [2] A. Chetouani, A. Beghdadi and M. Deriche. A new no reference image quality index for

Ringin : La mesure de ringin utilisée ici est celle proposée par [SBdV05]^[3]. Rappelons qu'elle est basée sur une phase d'apprentissage dans le domaine des ondelettes.

La base LIVE a été utilisée [SWCB06]^[4]. Ce choix est essentiellement lié au nombre important d'images dégradées disponibles (plus de 150 images par dégradation). Une description de cette base a été présentée dans le chapitre 1.

Ces métriques sont ensuite combinées de façon à obtenir un index de qualité unique. L'outil de fusion utilisé ici est un réseau de neurones de type MLP dont les caractéristiques sont les suivantes.

Entrées : 3 correspondantes aux métriques sans référence sélectionnées.

Couche cachée : 1 avec 19 neurones.

Sortie : 1 dont la cible est le MOS.

Une fonction d'activation : Fonction sigmoïde.

Apprentissage : Algorithme de rétro-propagation.

5.3.2 Résultats expérimentaux

L'efficacité de la méthode proposée a été évaluée en termes de corrélation avec l'appréciation subjective humaine. Les corrélations obtenues pour chacune des mesures sont d'abord présentées dans le tableau 5.6.

Le tableau 5.7 présente les corrélations obtenues par la méthode proposée pour chaque dégradation. Nous pouvons constater que les performances ont été

blur estimation in the frequency domain. *International Symposium on Signal Processing and Information Technology*, 2009.

- [3] H.R. Sheikh, A.C. Bovik and G. de Veciana. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, Vol. 14, pp. 2117-2128, 2004.
- [4] H.R. Sheikh, Z. Wang, L. Cormack and A.C. Bovik. LIVE Image Quality Assessment Database. <http://live.ece.utexas.edu/research/quality>, 2006.

Métrique sans référence	Corrélation de Pearson
Effets de blocs	0.8160
Flou	0.8673
Ringings	0.9108

TABLE 5.6 – Corrélations de Pearson obtenues par chaque métrique (avant fusion).

améliorées pour toutes les dégradations, excepté pour le flou où une légère diminution est notée. Cette diminution est sans doute liée au fait que le flou est aussi une dégradation inhérente aux images compressées par JPEG2000.

Métrique sans référence	Corrélation de Pearson
Effets de bloc	0.9570
Flou	0.8462
Ringings	0.9289

TABLE 5.7 – Corrélations de Pearson obtenues par la méthode proposée, pour chaque dégradation (après fusion).

Le tableau 5.8 présente les corrélations obtenues globalement (toutes dégradations confondues). Une corrélation de Pearson de 0.89 est obtenue.

Corrélation	Valeur
Pearson	0.89964
Spearman	0.86214

TABLE 5.8 – Corrélations obtenues par la métrique proposée.

Pour mieux visualiser les résultats obtenus, l'index de qualité NR-IQM est affiché pour un échantillon d'images tests (voir figure 5.11). Il est à noter que les valeurs de sorties décrivent bien la forme des valeurs de référence (MOS).

5.4 Conclusions

Dans ce chapitre, nous nous sommes focalisés sur l'élaboration d'un système de fusion original permettant d'exploiter au mieux les performances des métriques existantes. La figure 5.12 illustre les solutions proposées dans cette étude.

Mesure avec référence :

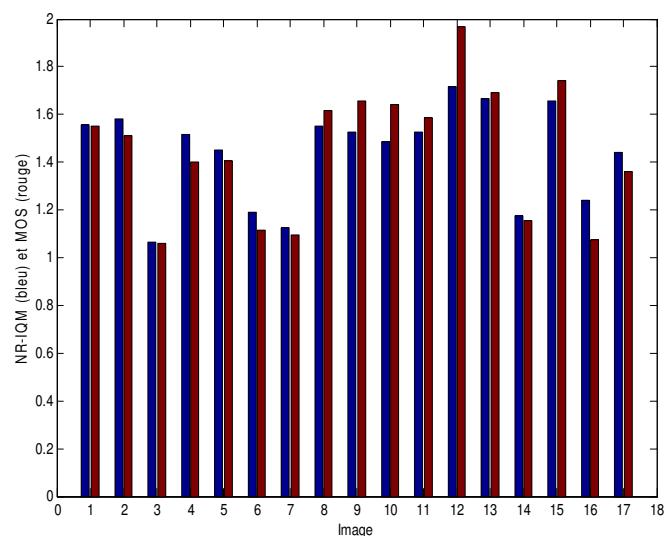


FIGURE 5.11 – Index de qualité pour différentes images : MOS (rouge) et NR-IQM (bleu).

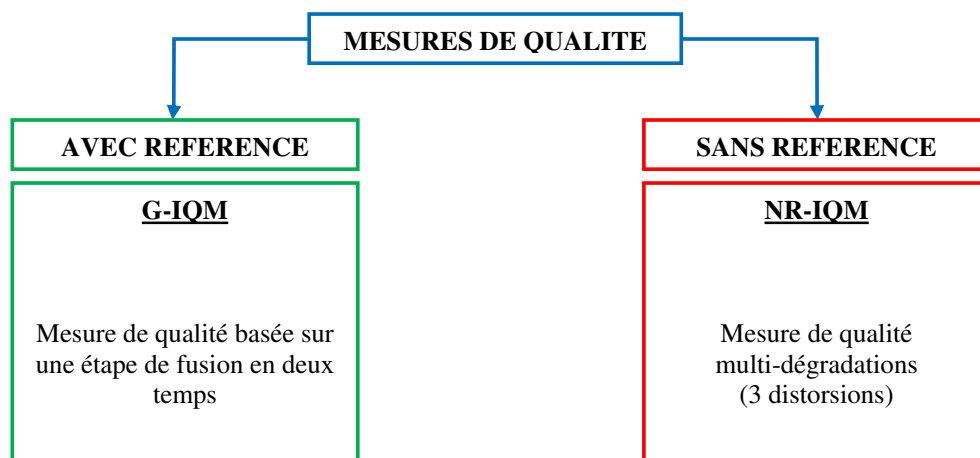


FIGURE 5.12 – Schéma récapitulatif.

Nous avons proposé une nouvelle approche permettant d'améliorer les performances des mesures existantes. Pour une dégradation donnée, les métriques sont d'abord classées selon le coefficient de corrélation de Pearson. Les trois meilleures sont ensuite sélectionnées et combinées de façon à obtenir une mesure unique par dégradation (IQM/D), permettant ainsi d'améliorer l'estimation de la qualité (restreinte à une dégradation donnée). La mesure glo-

bale est ensuite obtenue en combinant les différentes métriques IQM/Ds. La mesure proposée a été testée et validée sur un ensemble conséquent d'images dégradées.

Mesure sans référence :

Nous avons ensuite proposé une métrique de qualité sans référence multi-dégradations, permettant de dépasser une des hypothèses implicite, intrinsèque dans les approches classiques où l'on suppose la dégradation comme à priori connue. Différentes mesures ont été sélectionnées et combinées pour obtenir un index de qualité unique. Les tests expérimentaux présentés ont permis de valider l'approche proposée.

Chapitre 6

Système de mesure complet

Sommaire

6.1	Mesure avec référence : Système complet	162
6.1.1	Caractérisation des distorsions	162
6.1.2	Classifieur	164
6.1.3	Résultats expérimentaux	165
6.2	Mesure sans référence : Système complet	168
6.2.1	Caractérisation des distorsions	169
6.2.1.1	Effets de bloc	169
6.2.1.2	Flou	170
6.2.1.3	Ringings	171
6.2.2	Classifieur	171
6.2.3	Résultats expérimentaux	171
6.3	Conclusions	174

Le chapitre précédent a permis de mettre en évidence la faiblesse des métriques de qualité avec et sans référence. Nous avons ainsi proposé des solutions permettant de pallier ces insuffisances.

Nous avons constaté que les performances des métriques (avec référence) pour chaque dégradation sont relativement meilleures (IQM/D). L'étape de fusion globale a ainsi engendré une atténuation des performances. Il semble donc judicieux d'essayer de s'affranchir de cette étape de fusion.

L'objectif ici est alors de proposer une méthode permettant de détecter le type de dégradation contenu dans l'image afin de sélectionner la métrique IQM/Ds la plus appropriée. Cette approche est appliquée aux mesures avec et sans référence. Il est à noter que pour les mesures sans référence, l'objectif est ici différent. Le but est de détecter le type de distorsion pour s'affranchir de toute information a priori.

Dans ce qui suit, nous présentons d'abord le système complet de mesure de qualité d'image avec référence proposé. Puis, nous décrivons le système complet d'estimation de qualité d'image sans référence.

6.1 Mesure avec référence : Système complet

Dans ce système, la première phase repose sur l'extraction des descripteurs de l'image originale et de sa version dégradée pour chacune des dégradations considérées. Les descripteurs sont projetés dans un nouvel espace, permettant de mieux séparer les distorsions. Le type de dégradation peut être ainsi identifié selon un critère de distance. La figure 6.1 présente le schéma synoptique de la méthode proposée. La base TID 2008 est utilisée dans le cadre de cette étude.

Dans ce qui suit, nous présentons d'abord la nature des descripteurs extraits ainsi que le classifieur employé. Le système complet est ensuite décrit.

6.1.1 Caractérisation des distorsions

Pour caractériser chacune des dégradations considérées, un ensemble de descripteurs est extrait des images. Il s'agit de se placer dans un espace mutlidimensionnel

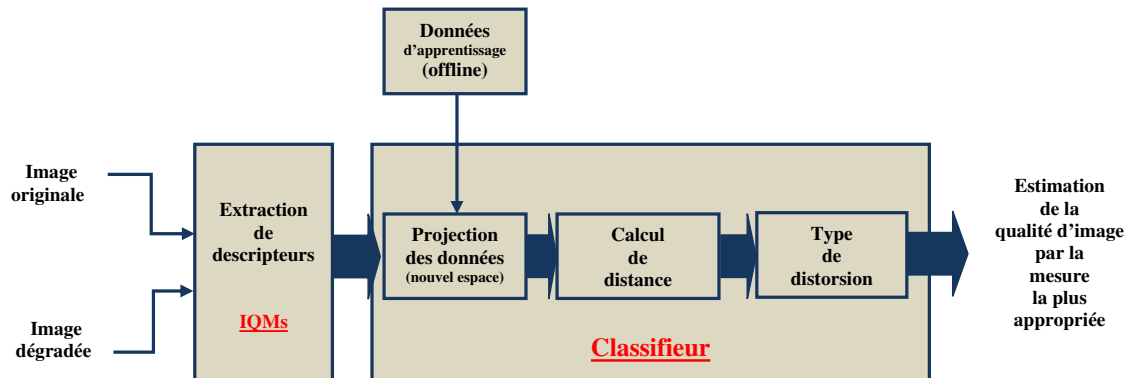


FIGURE 6.1 – Schéma synoptique.

dans lequel les distorsions considérées et les mesures de qualité sont les éléments constitutants. Nous proposons ici d'utiliser directement les métriques de qualité pour caractériser les dégradations et les détecter. L'idée sous-jacente à ce raisonnement repose sur le fait que la qualité d'image est un problème multidimensionnel. Cela a déjà été noté par Daly et Boohart [JAC93]^[1] [Mar02]^[2].

Dans ce travail, nous nous sommes limités aux métriques de qualité décrites dans le chapitre précédent, à savoir : PSNR, UQI [WB02b]^[3], SSSIM et MSSIM

-
- [1] A. J. Ahumada, Jr. and C. H. Null. Image quality: A multidimensional problem. *IEEE Signal Processing Magazine*, Vol. 18, pp. 851-854, 1992.
 - [2] J-B. Martens. Multidimensional modeling of image quality. *Proceedings of IEEE*, Vol. 90, pp. 133-153, 2002.
 - [3] Z. Wang and C. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, Vol. 2, pp. 81-84, 2002.

[WBE05]^[1], PSNR-HVS et PSNR-HVSM [PSE⁺07]^[2] [PCL⁺08]^[3], WSNR [MV93]^[4], IFC [SBdV05]^[5], VIF et VIFP [SB06a]^[6]. Il est à noter que d'autres mesures peuvent être utilisées au prix de l'accroissement de la dimension de l'espace.

6.1.2 Classifieur

Une fois les descripteurs choisis et extraits, il s'agit maintenant de les utiliser dans un processus d'identification des dégradations. Nous proposons ici d'utiliser l'Analyse Discriminante Linéaire (LDA) comme outil de classification des distorsions. La LDA est une méthode populaire de réduction de dimension et de classification dont le principe de base repose sur la projection des données dans un espace permettant un niveau de séparabilité maximum. Cet espace de projection est déterminé par minimisation de la variance intra-classe et la maximisation de la variance inter-classe [PBHK97]^[7]. Les classes correspondent ici aux dégradations.

Ainsi, pour chaque distorsion, la base d'images est subdivisée en deux sous

-
- [1] Z. Wang, A.C. Bovik and E.P. Simoncelli. Structural approaches to image quality assessment. *Handbook of image and video processing, 2nd Edition, Al Bovik, ed, Academic Press*, 2005.
 - [2] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola and V. Lukin. On between-coefficient contrast masking of DCT basis functions. *Workshop on Video Processing and Quality Metrics*, 2007.
 - [3] N. Ponomarenko, M. Carli, V. Lukin, K. Egiazarian, J. Astola and F. Battisti Color Image Database for Evaluation of Image Quality Metrics. *Proceedings of International Workshop on Multimedia Signal Processing*, pp. 403-408, 2008.
 - [4] T. Mitsa and K. Varkur Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms. *International Conference on Acoustics, Speech, and Signal Processing*, pp. 301-304, 1993.
 - [5] H.R. Sheikh, A.C. Bovik and G. de Veciana An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, Vol. 14, pp. 2117-2128, 2004.
 - [6] H.R. Sheikh and A.C. Bovik Image information and visual quality. *IEEE Transactions on Image Processing*, Vol. 15, pp. 430-444, 2006.
 - [7] N. Peter, J. Belhumeur, P. Hespanha and D. Kriegman. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 711-720, 1997.

ensembles :

- **Ensemble d'apprentissage** : utilisé uniquement lors de la phase d'apprentissage (60%).
- **Ensemble de tests** : utilisé uniquement pour tester l'efficacité de la méthode proposée (40%).

6.1.3 Résultats expérimentaux

Pour l'évaluation des performances, un panel de plus de 400 images différentes de la base d'apprentissage est utilisé. La procédure de test est simple et requiert l'image originale et sa version dégradée.

La figure 6.2 illustre un exemple de fonctionnement pour une image compressée par JPEG. La procédure est la suivante :

1. Extraction des descripteurs (index de qualité).
2. Projection des descripteurs dans le nouvel espace (défini durant la phase d'apprentissage).
3. Comparaison des descripteurs projetés dans le nouvel espace et les descripteurs de chaque classe.
4. Détection du type de dégradation selon le critère de distance minimale.

Dans notre étude, la distance euclidienne s'est révélée efficace. Nous avons aussi testé la distance de Mahalanobis, les performances obtenues sont équivalentes mais avec un temps de calcul supérieur.

L'efficacité du système proposé a été évaluée en termes de bonne classification. La figure 6.3 présente le pourcentage de bonne classification pour chaque type de dégradation. Nous pouvons noter que ce pourcentage est toujours supérieur à 90% pour tous les artefacts considérés. Il est égal en moyenne à 98.11%.

Pour mettre en évidence les erreurs de classification, la matrice de confusion est calculée (voir tableau 6.1). Les taux de classification les plus faibles sont obtenus

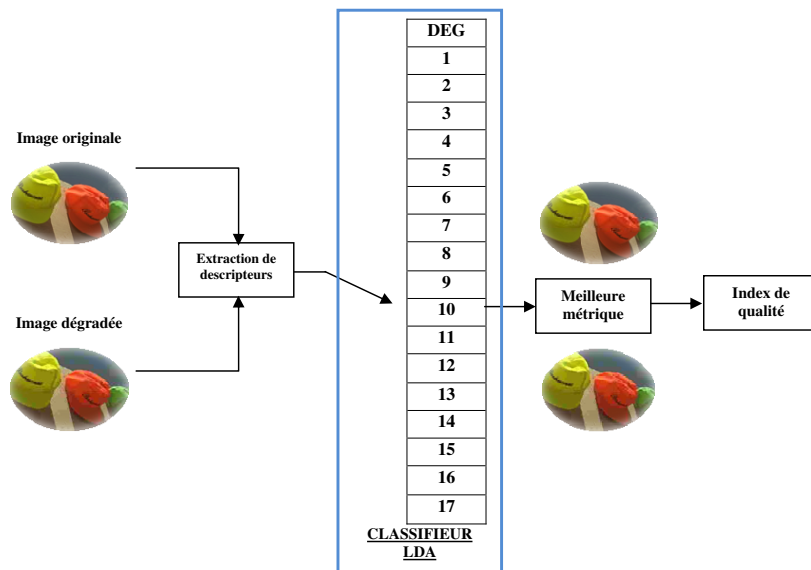


FIGURE 6.2 – Schéma fonctionnel (exemple de distorsion JPEG).

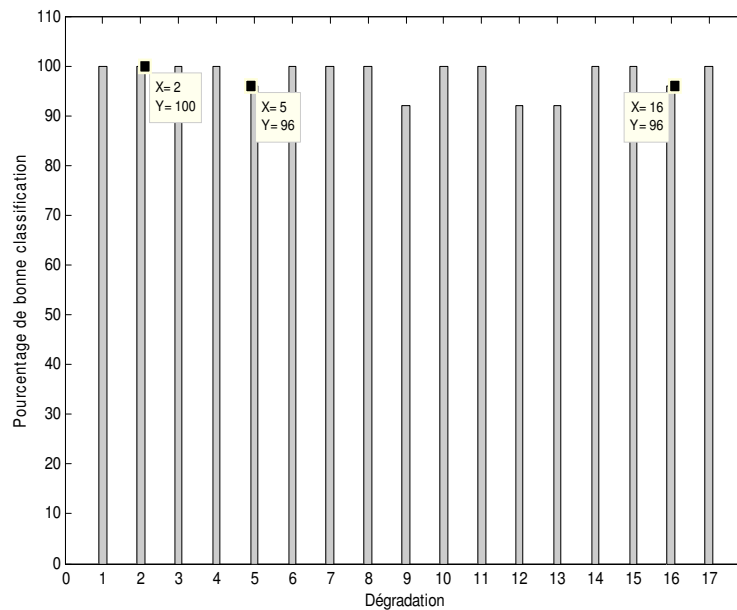


FIGURE 6.3 – Pourcentage de bonne classification.

pour les classes 9, 12 et 13 (92%) qui correspondent aux dégradations : lissage, erreur de transmission JPEG et erreur de transmission JPEG2000.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	4	96	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	92	4	4	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	92	8	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	8	92	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
16	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	96	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100

TABLE 6.1 – Matrice de confusion.

Ces confusions sont généralement dues à l'apparition de plusieurs dégradations dans la même image, lors de fort taux de compression, comme c'est le cas notamment des images compressées par JPEG2000 (ringing et flou).

Un échantillon d'images pour lesquelles la confusion a lieu est présenté par la figure 6.4 (confusion de la dégradation appelée lissage (8) avec les distorsions de type JPEG (10) et JPEG2000 (11)). On s'aperçoit que les distorsions de l'image dégradée de gauche ressemblent aux artefacts engendrés par JPEG2000 (ringing et flou), tandis que l'image dégradée de droite contient des dégradations visuellement similaires aux effets de bloc.

Le système d'estimation de qualité d'image complet se présente finalement sous la forme suivante (voir figure 6.5). Un tel système permet ainsi d'estimer la nature et le niveau des distorsions de l'image.



FIGURE 6.4 – Exemple d’erreurs de classification : dégradation 8 confondue avec les artefacts 11 (gauche) et 10 (droite).

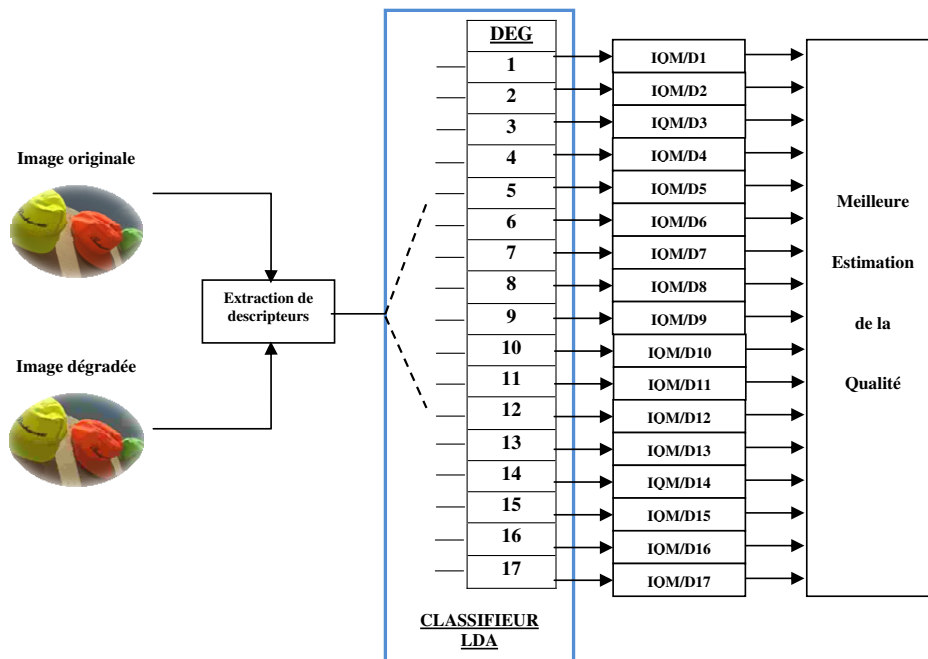


FIGURE 6.5 – Classification de dégradation : Système complet.

6.2 Mesure sans référence : Système complet

Nous proposons maintenant de reprendre le principe précédent et de l’appliquer à l’estimation de la qualité d’image sans référence. La figure 6.6 présente le schéma synoptique du système proposé. La première étape consiste à extraire les descripteurs correspondant aux métriques de qualité dédiées à quelques dégradations les

plus connues. Le type de distorsion est ensuite déterminé alors par la sortie du classifieur.

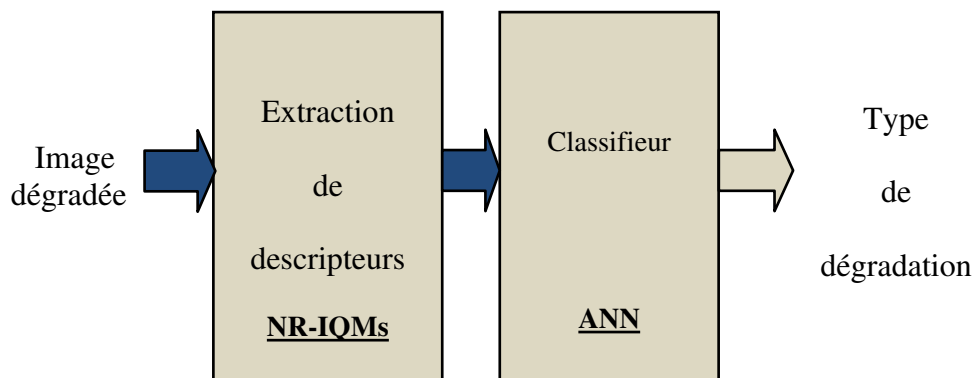


FIGURE 6.6 – Schéma synoptique.

Dans ce qui suit, les descripteurs ainsi que le classifieur utilisé sont dans un premier temps présentés. Les résultats expérimentaux sont ensuite exposés et discutés. Nous nous limitons dans cette étude aux dégradations les plus fréquentes à savoir les effets de blocs, le flou et le ringing.

6.2.1 Caractérisation des distorsions

Comme pour les mesures avec référence, les descripteurs choisis ici sont les mesures de qualité sans référence. Nous avons utilisé au total 5 métriques : 3 pour le flou, 1 pour les effets de blocs et 1 pour le ringing. Chacune de ces mesures est brièvement décrite dans cette section.

6.2.1.1 Effets de bloc

La mesure des effets de blocs utilisée ici est basée sur une analyse fréquentielle de l'image [WBE00]^[1]. Les auteurs proposent de détecter des pics à certaines

[1] Z. Wang, A.C. Bovik and B.L. Evans. Blind measurement of blocking artefacts in images. *International Conference on Image Processing*, 2000.

fréquences particulières qui correspondent au partitionnement de l'image en blocs 8x8. L'amplitude des pics est utilisée comme mesure.

6.2.1.2 Flou

Trois mesures de flou sans référence ont ici été utilisées. La première mesure est basée sur la transformée en ondelettes [TLZZ04]^[1]. Pour chaque niveau de décomposition, une carte de contours est calculée à partir des coefficients hautes fréquences. L'index de qualité est ensuite déterminé via une analyse du type de contour contenu dans l'image.

La deuxième mesure utilisée a été proposée par [Crê07]^[2]. L'idée principale développée par les auteurs est qu'une image floue est peu sensible à un ajout de flou supplémentaire. A partir de cette observation, les auteurs proposent d'estimer le flou en analysant l'impact de la distorsion ajoutée. La mesure globale est obtenue par analyse des variations locales de l'image et de sa version floue.

Nous avons aussi utilisé la métrique décrite dans le chapitre 3 [CBD09]^[3]. Celle-ci est basée sur une analyse radiale du spectre de puissance inspirée de la mesure précédente [Crê07]^[2]. L'impact du flou ajouté est ici mesuré dans le domaine fréquentiel.

-
- [1] H. Tong, M. Li, Zhang and C. Zhang. Blur detection for digital images using wavelet transform. *International Conference on Multimedia and Expo, Vol. 1, pp. 17-20*, 2004.
 - [2] F. Crête. Estimer, mesurer et corriger les artefacts de compression pour la télévision. *Rapport de Thèse, Université Joseph Fourier*, 2007.
 - [3] A. Chetouani, A. Beghdadi and M. Deriche. A new no reference image quality index for blur estimation in the frequency domain. *International Symposium on Signal Processing and Information Technology*, 2009.

6.2.1.3 Ringing

Comme mesure de ringing, la métrique proposée par [SBC05]^[4] a été ici utilisée. Elle est basée sur un apprentissage dans le domaine des ondelettes. Une fois la transformée appliquée, une phase d'apprentissage permet d'estimer différents paramètres et seuils. Ces paramètres sont ensuite utilisés pour estimer la dégradation.

6.2.2 Classifieur

Un grand nombre de méthodes de classification de données a été proposé dans la littérature. Dans cette étude, nous avons opté pour l'approche neuronale. Néanmoins, nous avons aussi testé la méthode basée sur la LDA, de meilleures performances ont été obtenues avec les réseaux de neurones. La structure du modèle final, présentée par la figure 6.7, est définie comme suit :

- **Entrées** : 5 (i.e. mesures sans référence).
- **Couche cachée** : 1 (le nombre de neurones est égal à 9).
- **Sorties** : 3 (i.e. nombre de dégradations).
- **Fonction d'activation** : Sigmoidale.
- **Apprentissage** : Algorithme de rétro-propagation.

6.2.3 Résultats expérimentaux

La méthode proposée a été évaluée en termes de taux de bonne classification. La base TID 2008 a été ici utilisée. L'ensemble de la base a été subdivisé en deux sous ensembles : un ensemble d'apprentissage et un ensemble test. Un peu moins de 200 images naturelles différentes de celles de l'ensemble d'apprentissage ont été utilisées pour les tests. La figure 6.8 résume le principe de fonctionnement du système.

Le processus est simple et ne requiert que l'image dégradée. Pour une image

[4] H. R. Sheikh, A. C. Bovik and L. K. Cormack. No-Reference Quality Assessment Using Natural Scene Statistics: JPEG2000. *IEEE Transactions on Image Processing*, Vol. 14, 2005.

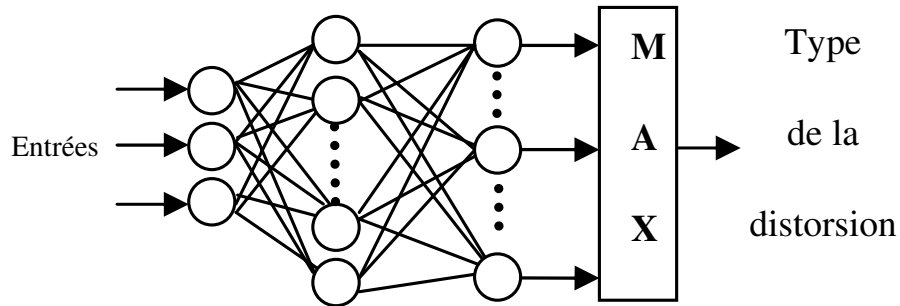


FIGURE 6.7 – Classifieur utilisé.

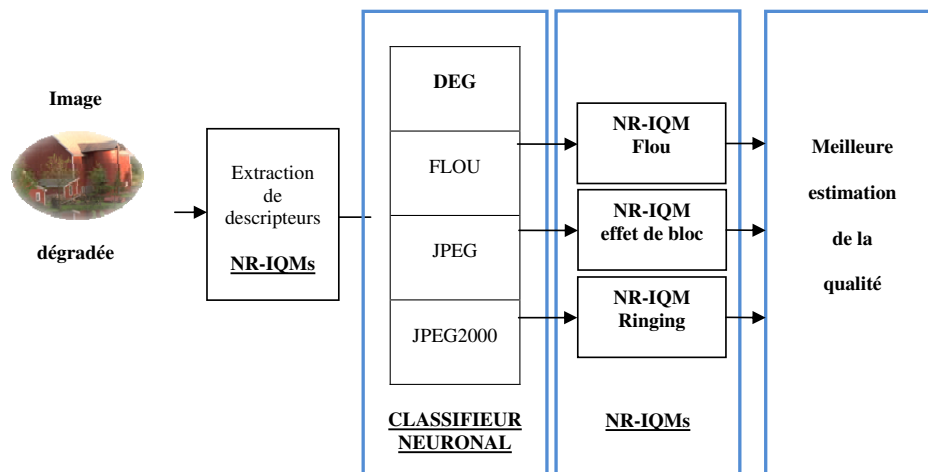


FIGURE 6.8 – Système de mesure sans référence complet.

donnée, les descripteurs sont d'abord extraits et envoyés à l'entrée du réseau de neurones. La valeur maximale des sorties du classifieur permet de déterminer ensuite le type de dégradation contenu dans l'image.

Le pourcentage moyen de bonne classification obtenu est de 96.55%. La figure 6.9 présente le pourcentage obtenu pour chacune des distorsions considérées.

Afin de mieux visualiser les erreurs de classification, la matrice de confusion

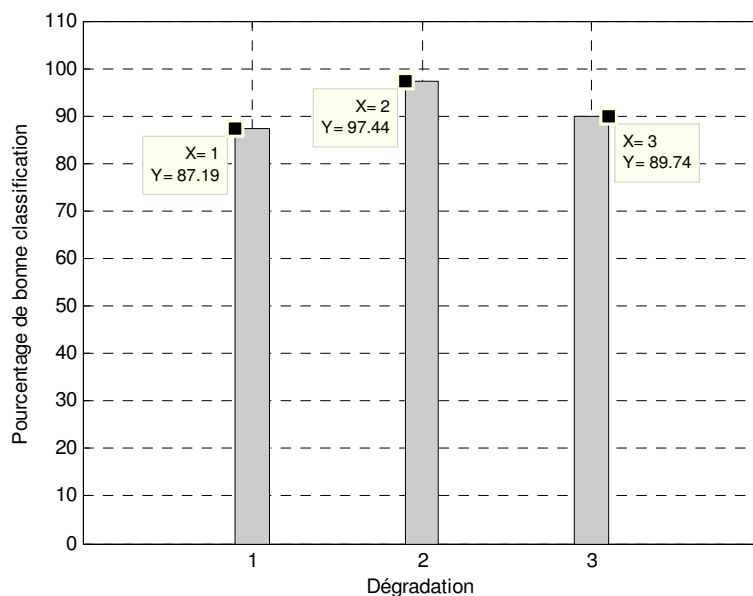


FIGURE 6.9 – Pourcentage de bonne classification.

est calculée et présentée par le tableau 6.2.

	Flou	JPEG	JPEG2000
Flou	87.179	0	12.82
JPEG	0	97.43	2.5
JPEG2000	10.25	0	89.74

TABLE 6.2 – Matrice de confusion.

Pour le flou et JPEG2000, il est à noter que les confusions sont les plus élevées. Cela est essentiellement dû au fait que le flou est une dégradation inhérente de la compression JPEG2000 (6.10). En effet, nous pouvons voir que la dégradation principale qui apparaît dans les images compressées par JPEG2000 est l'effet de "ringing". Cependant, certaines zones de l'image contiennent aussi du flou.

Nous affichons dans la figure 6.11 les images pour lesquelles il y a confusion. Ces erreurs de classification apparaissent généralement pour des images fortement compressées.

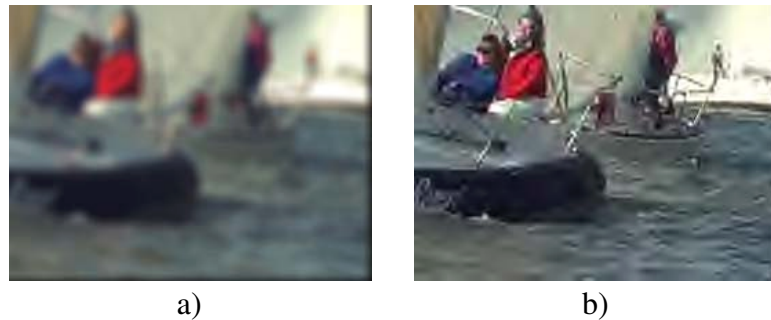


FIGURE 6.10 – a) Image flou et b) compressée par JPEG2000.

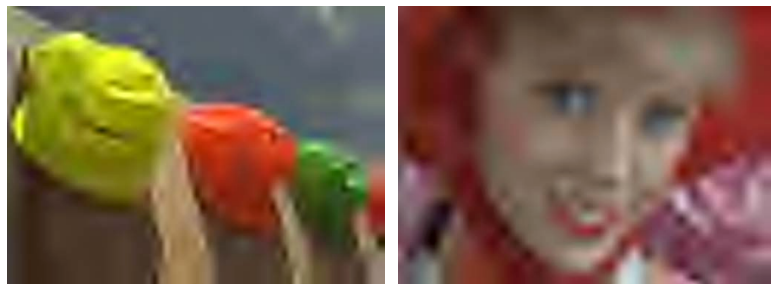


FIGURE 6.11 – Exemple d'erreur de confusion : classe flou confondue avec la classe JPEG2000.

6.3 Conclusions

Dans ce chapitre, nous avons proposé un système d'estimation de la qualité d'image complet permettant de détecter automatiquement le type de distorsion contenu dans une image pour ensuite utiliser la métrique la plus appropriée. La représentation globale du système est illustrée par la figure 6.12.

Mesure avec référence : système complet

Nous avons proposé une nouvelle approche efficace permettant d'estimer la qualité d'une image par la métrique la plus appropriée. Une phase de détection du type de dégradation contenu dans l'image permet de sélectionner la mesure la plus adaptée à la dégradation détectée. L'étape de détection est ici réalisée par une analyse discriminante linéaire dont les descripteurs sont des index de qualité. La méthode proposée a été validée sur une base assez

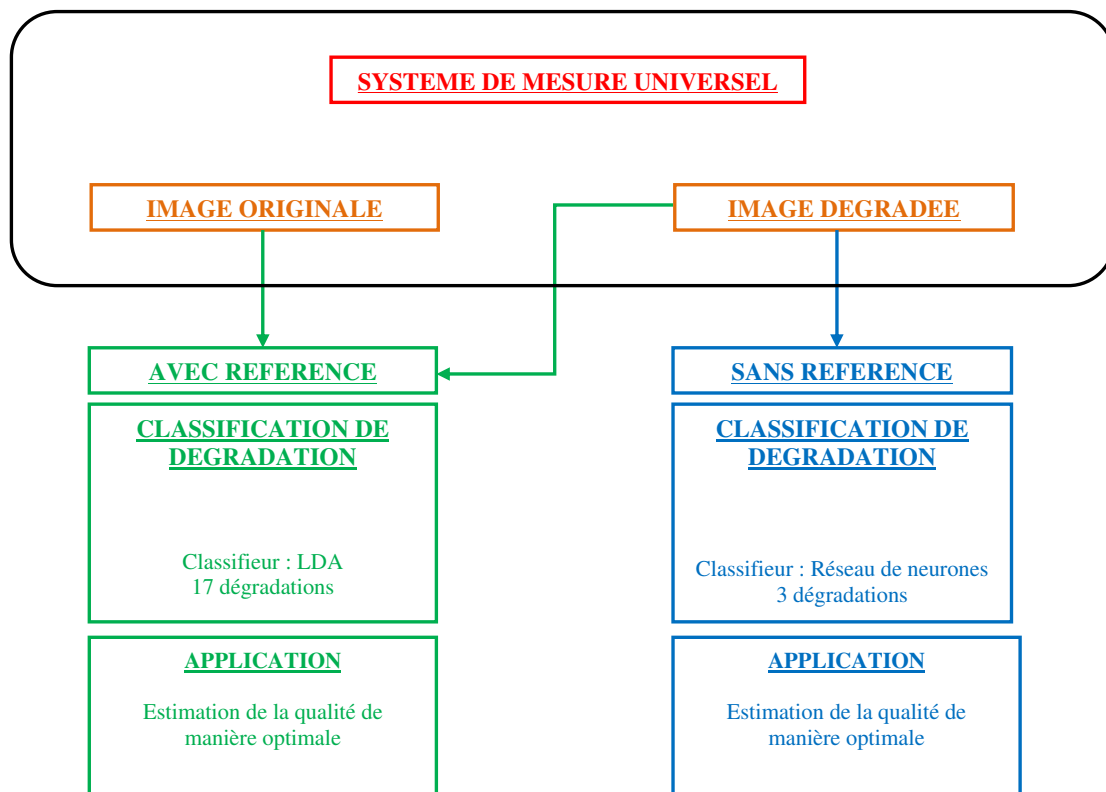


FIGURE 6.12 – Système de mesure universel.

riche.

Mesure sans référence : système complet

Le principe de classification de dégradation avec référence a été repris et appliqué aux mesures sans référence. Le nombre de distorsions considéré s'élève ici à 3. Cette méthode permet ainsi de s'affranchir de toute connaissance a priori et particulièrement celle liée au type d'artefact. Les résultats obtenus sont encourageants et confirment l'efficacité de la méthode proposée.

Conclusion générale

Nous avons à travers ce travail abordé plusieurs aspects liés à la qualité d'image suivant une démarche incrémentale et progressive. Dans un premier temps, nous nous sommes intéressés à la mesure et prédiction de quelques dégradations. Nous nous sommes limités aux trois artefacts les plus connus et en général les plus gênants : le flou, les effets de bloc et l'effet de Gibbs 2D ou "ringing". Pour chacune de ces distorsions nous avons proposé des métriques qui se sont révélées très corrélées à l'appréciation subjective.

Compte tenu du caractère multidimensionnel de la qualité d'image, nous avons proposé une approche coopérative où l'on combine les mesures les plus adaptés aux artefacts étudiés suivant un processus d'apprentissage. La prise en compte des trois aspects : identification, estimation et décision s'inscrit dans une démarche très inspirée du Système Visuel Humain. Nous pensons aussi, comme la plupart de nos collègues travaillant dans le domaine, que la notion de qualité d'image est avant tout subjective et qu'il convient d'intégrer l'observateur au centre de nos réflexions. Cependant, nous sommes persuadés qu'une approche totalement inspirée des mécanismes du SVH est vouée à l'échec compte tenu des connaissances limitées et de la nature imprédictible du comportement humain. Nous croyons à une approche où l'on combine de façon mesurée les aspects liés directement au SVH et les techniques de traitement d'images.

Nous avons à travers cette étude montré qu'il est possible de converger vers un système complet d'évaluation objective de la qualité d'image le plus corrélé possible avec l'appréciation subjective. Il reste certes d'autres aspects à approfondir et à améliorer et notamment la prise en compte d'autres distorsions et le choix de méthodes et de modèles d'apprentissage plus avancés. Nous avons montré

qu'avec des modèles classiques de classification et d'apprentissage, qu'il était possible de quantifier de façon fiable trois des dégradations les plus gênantes et les plus répandues dans les systèmes de capture et codage d'images.

Nous proposons comme perspectives de recherche, de poursuivre ce travail dans les domaines suivants :

- **Effet de bloc** : Il serait intéressant d'exploiter la carte de prédiction des effets de bloc dans d'autres applications et de l'intégrer notamment dans un schéma de compression bloc avec contrôle de qualité. Des tests ont été menés dans ce sens, les résultats préliminaires semblent convaincants et restent à valider.
- **Flou et ringing** : Les méthodes d'estimation du flou et de ringing proposées peuvent être intégrées dans un schéma de réduction de ces artefacts. Des tests ont été réalisés, les résultats obtenus semblent confirmer l'approche proposée et restent à valider dans des bases d'images variées.
- **Autres artefacts** : Nous envisageons d'enrichir le système de mesure complet par la prise en compte d'autres distorsions. Nous avons montré qu'un tel système est possible. Dans l'état actuel de nos recherches, nous ne prétendons pas avoir fourni une solution complète. La prise en compte d'autres dégradations et particulièrement celles liées à la couleur est un axe de recherche à développer.
- **Prise en compte des cartes de saillance** : l'exploitation des cartes de saillances perceptuelles est une poursuite naturelle de nos investigations dans la mise au point d'un système le plus complet possible.
- **Extension à la vidéo** : les aspects spatio-temporels nécessaires à l'analyse de la vidéo n'ont pas été considérés dans ce travail. C'est une des pistes à explorer dans un avenir proche. Nous avons déjà commencé à aborder cette question dans le cadre d'un projet de recherche appliquée.

Bibliographie

- [A.B87] A.B.Watson. The cortex transform : Rapid computation of simulated neural images. *Computer Vision Graphics and Image Processing*, 39 :311–327, 1987.
- [AMPD91] J-P. Antoine, R. Murenzi, B. Piette, and M. Duval. Image analysis with 2d continuous wavelet transform : Detection of position, orientation and visual contrast of simple objects. *Wavelets and Applications*, Y. Meyer Editor, Masson/Springer-Verlag, 11 :144–159, 1991.
- [ARM04] G. Medeghini A. Rizzi, T. Algeri and D. Marini. A proposal for contrast measure in digital images. *CGIV04, IST Second European Conference on Color in Graphics, Imaging and Vision*, pages 187–192, 2004.
- [BA83] P. J. Burt and H. E. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communication*, 31 :532–540, 1983.
- [Bar90] P.G.J. Barten. Evaluation of subjective image quality with the square-root integral method. *Journal of Optical Society of America*, 7 :2024–2031, 1990.
- [BCFM02] D. Bailey, M. Carli, M. Farias, and S. Mitra. Quality assessment for block-based compressed images and videos with regard to blockiness artifacts. *International Workshop in Data Compression*, 2002.
- [BD00] A. Beghdadi and M. Deriche. Features extraction from fingerprints using frequency analysis. *Workshop on Signal Processing and Applications*, 18 :14–15, 2000.

- [Bek99] N. BekkatAT. Critère objectifs de qualité subjective d'images monochromes. conception du modèle et validation expérimentale, rapport de thèse, université de nantes - ireste, 1999.
- [BHB04] A. Bouzerdoun, A. Havstad, and A. Beghdadi. Image quality assessment using neural networks approach. *International Joint Conference on Neural Networks*, pages 330–333, 2004.
- [BI01] A. Beghdadi and R. Iordache. A wigner-ville distribution-based image dissimilarity measure. *International Symposium on Signal Processing and Its Applications*, 2 :430–433, 2001.
- [BL01] A.C. Bovik and S. Liu. Dct-domain blind measurement of blocking artefacts in dct-coded images. *International Conference on Acoustics, Speech, and Signal Processing*, pages 1725–1728, 2001.
- [Bli03] J.L. Blin. Samviq - subjective assessment methodology for video quality. *rapport technique BPN 056, EBU Project Group B/VIM Video Multimedia*, 2003.
- [BM98] M. Bolin and G. Meyer. Image quality assessment by using neural networks. *Special Interest Group on Computer Graphics and Interactive Techniques Conference*, 5 :409–418, 1998.
- [BM99] M. Bolin and G. Meyer. A visual difference metric for realistic image synthesis. *SPIE, Human Vision and Electronic Imaging IV*, 3644 :106–120, 1999.
- [BN89] A. Beghdadi and A. Le Négrate. Contrast enhancement technique based on local detection of edges. *Computer Vision, Graphics, and Image Processing*, pages 162–174, 1989.
- [BP05] R.V. Babu and A. Perkis. An hvs-based no reference perceptual quality assessment of jpeg coded images using neural networks. *International Conference on Image Processing*, pages 433–439, 2005.
- [BP06] A.C. Brooks and T.N. Papas. Structural similarity quality metrics in a coding context : exploring the space of realistic distortions. *SPIE, Human Vision and Electronic Imaging XI*, 6057 :299–310, 2006.
- [BPP03] A. Beghdadi and B. Pesquet-Popescu. A new image distortion measure based wavelet decomposition. *International Symposium on Signal Processing and Its Applications*, 1 :485–488, 2003.

- [Bra99] A. P. Bradley. A wavelet visible difference predictor. *IEEE Transactions on Image Processing*, 8 :717–730, 1999.
- [BS05] R. Barland and A. Saadane. reference free quality metric for jpeg2000. *International Symposium on Signal Processing and Its Applications*, 1 :351–354, 2005.
- [BT.02] BT.500-11. Methodology for the subjective assessment of the quality of television pictures. *International Telecommunication Union*, 2002.
- [CA05] P. Le Callet and F. Autrusseau. Subjective quality assessment ircyn/ivc database. <http://www.irccyn.ec-nantes.fr/ivcdb/>, 2005.
- [CBD09] A. Chetouani, A. Beghdadi, and M. Deriche. A new no reference image quality index for blur estimation in the frequency domain. *International Symposium on Signal Processing and Information Technology*, 2009.
- [CGC98] F.X. Coudoux, M.G. Gazalet, and P. Corlay. Reduction of blocking effect in dct-coded images based on a visual perception criterion. *Signal Processing : Image Communication*, 11 :179–186, 1998.
- [CH07a] D. Chandler and S. Hemami. Subjective image database. <http://fowlard.ece.cornell.edu/dmc27/vsnr/vsnr.html>, 2007.
- [CH07b] D.M. Chandler and S. Hemami. Vsnr : A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing*, 16 :2284–2298, 2007.
- [Cha10] D. Chandler. Content-based strategies of image and video quality assessment. <http://vision.okstate.edu/index.php?loc=csiq>, 2010.
- [CHG00] S. Chen, Z. He, and P.M. Grant. Artificial neural network visual model for image quality enhancement. *Neurocomputing*, 30 :339–346, 2000.
- [CHGZ02] P. Carrai, I. Heynderickx, P. Gastaldo, and R. Zunino. Image quality assessment by using neural networks. *International Symposium on Circuits and Systems*, 5 :253–256, 2002.
- [CJ01] J. Caviedes and J. Jung. No-reference metric for a video quality control loop. *International Conference on Information Systems, Analysis and Synthesis*, pages 290–295, 2001.

- [CL95] C.H. Chou and Y.C. Li. A perceptually tuned subband image coder based on the measure of just-noticeable distortion profile. *IEEE Transactions on Circuits and Systems for Video Techniques*, 5 :467–476, 1995.
- [CL05] H. Cheng and J. Lubin. Reference free objective quality metrics for mpeg coded video. *SPIE, Human vision and electronic imaging X*, 5666 :160–167, 2005.
- [CLS05] C. Charrier, C. Larabi, and H. Saadane. Evaluation de la qualité des images. *Ecole d'hiver sur l'image numérique couleur*, 2005.
- [CMB08] A. Chetouani, G. Mostafaoui, and A. Beghdadi. Predicting blocking effects in the spatial domain using a learning approach. *International Conference on Signal Processing and Multimedia Application*, pages 197–201, 2008.
- [CO04] J. Caviedes and F. Oberti. A new sharpness metric based on local kurtosis, edge and energy information. *Signal Processing : Image Communication*, 19 :147–163, 2004.
- [Cor70] T. N. Cornsweet. Visual perception. *Academic Press, New York*, 1970.
- [Crê07] F. Crête. Estimer, mesurer et corriger les artefacts de compression pour la télévision. *Manuscrit de Thèse, Université Joseph Fourier*, 2007.
- [CSB99] P. Le Callet, A. Saadane, and D. Barba. Orientation selectivity of opponent-colour channels. *European Conference on Visual Perception*, 28 :67, 1999.
- [CYA06] G. Chen, C.L. Yang, and S.L. Andxie. Gradient-based structural similarity for image quality assessment. *International Conference on Image Processing*, pages 2929–2932, 2006.
- [CYPX06] G. Chen, C. Yang, L. Po, and S. Xie. Edge-based structural similarity for image quality assessment. *International Conference in Acoustics, Speech and Signal Processing*, 2 :933–936, 2006.
- [Dal93] S. Daly. The visible differences predictor : an algorithm for the assessment of image fidelity. *Digital images and human vision*, pages 179–206, 1993.

- [Der87] R. Deriche. Using canny's criteria to derive a recursively implemented optimal edge detector. *International Journal of Vision*, 1 :167–187, 1987.
- [DTZY08] W. Ding, Y. Tong, Q. Zhang, and D. Yang. Image and video quality assessment using neural network and svm. *Tsinghua Science Technology*, 13(1) :112 – 116, 2008.
- [Dur] Durant. Méthodes d'évaluation subjective de la qualité audiovisuelles pour applications multimédias, recommandation uit-t p.911.
- [DVKG⁺00] N. Damera-Venkata, T. Kite, W. Geisler, B. Evans, and A.C. Bovik. Image quality assessment based on a degradation model. *IEEE Transactions on Image Processing*, 9 :636–650, 2000.
- [DXH07] Y. Dongxue, H. Xinsheng, and T. Hongli. Image quality assessment based on wavelet coefficients using neural network. *International Symposium on Neural Networks*, pages 853–859, 2007.
- [EAP⁺06] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli. New full-reference quality metrics based on hvs. *International Workshop on Video Processing and Quality Metrics*, 2006.
- [Far04] M.C.Q. Farias. No-reference and reduced reference video quality metrics : New contributions. *Ph.D Dissertation, Dept. of Electrical and Computer Engineering, University of California, Santa Barbara*, 2004.
- [FCNM04] M.C.Q. Farias, M. Carli, A. Neri, and S.K. Mitra. A new no-reference method for video quality assessment using data hiding. *Digital Signal Processing and its Applications*, 2004.
- [FK06] R. Ferzli and L. J. Karam. A human visual system-based no-reference objective image sharpness metric. *International Conference on Image Processing*, pages 2949–2952, 2006.
- [FK09] R. Ferzli and L.J. Karam. A no reference objective image sharpness metric based on the notion of jnb. *IEEE Transactions on Image Processing*, 18 :717–728, 2009.
- [Fra90] P. Franti. Les mécanismes de la vision. *Collection pour la science, Diffusion Belin*, 1990.

- [Fra98] P. Franti. Blockwise distortion measure for statistical and structural errors in digital images. *Signal Processing : Image Communication*, 13 :89–98, 1998.
- [FT99] E. Favier and A. Trémeau. Une mesure de contraste couleur inter-régions. *Groupe de Recherche et d'Etudes du Traitement du Signal et des Images*, pages 643–646, 1999.
- [GMUE02] D. Gayle, H. Mahlab, Y. Ucar, and A. M. Eस्कicioglu. A full-reference color image quality measure in the dwt domain. *European Signal Processing Conference*, 2002.
- [GR84] R. Gordon and R. M. Rangayan. Feature enhancement of film mammograms using fixed and adaptive neighborhood. *Journal of the Optical Society of America*, 23 :560–564, 1984.
- [GZVH02] P. Gastaldo, R. Zunino, E. Vicario, and I. Heynderickx. Cbp neural network for objective assessment of image quality. *International Joint Conference on Neural Networks*, 1 :194–199, 2002.
- [Hal81] C.H. Hall. Subjective evaluation of a perceptual quality metric. *SPIE - Image quality*, 310 :2000–2004, 1981.
- [Hou62] P.V.C Hough. Method and means for recognizing complex patterns. *U.S. Pattern*, 1962.
- [HT95] D.J. Heeger and P.C. Teo. A model of perceptual image fidelity. *International Conference on Image Processing*, 2, 1995.
- [htt] <http://www.hd3d.fr/en/>.
- [HWS10] R. Hassen, Z. Wang, and M. Salama. No reference image sharpness assessment based on local phase coherence measurement. *International Conference on Acoustics Speech and Signal Processing*, pages 2434–2437, 2010.
- [Ini08] Digital Cinema Initiatives. Digital cinema system specification. *LLC, Member Representatives Committee*, 2008.
- [ITU96] ITU. Itu recommendation p.930. *Principles of a reference impairment system for video*, 1996.
- [JAC93] J. Albert J.R. Ahumada and N.H. Cynthia. Image quality : a multi-dimensional problem. *Digital images and human vision, MIT Press*, pages 141–148, 1993.

- [Jol94] J.-M. Jolion. Analyse multirésolution du contraste dans les images numériques. *Traitement du Signal*, 11 :245–255, 1994.
- [KA90] F. Kammoun and J.P. Astruc. Augmentation de contraste adaptative suivant les niveaux de gris des contours. *Société française de physique*, 25 :1029–1035, 1990.
- [KB92] B. Kolpatzik and C. Bouman. Optimized error diffusion for high quality image display. *Journal Electronic Imaging*, pages 277–292, 1992.
- [KB95] B. W. Kolpatzik and C. A. Bouman. Optimized universal color palette design for error diffusion. *Journal Electronic Imaging*, 4 :131–143, 1995.
- [KH97] D. Kundur and D. Hatzinakos. A robust digital image watermarking method using wavelet-based fusion. *International Conference Image Processing*, 1 :544–547, 1997.
- [Kle93] S.A. Klein. Image quality and image compression : a psychophysicist's viewpoint. *Chapter in Digital Images and Human Vision*, ed. A.B. Watson, *PIT Press*, pages 73–88, 1993.
- [KP86] M. K. Kundu and S. K. Pal. Thresholding for edge detection using human psychovisual phenomena. *Pattern Recognition Letters*, 4 :433–441, 1986.
- [KPSP97] J. K. Kim, J. M. Park, K. S. Song, and H. W. Park. Adaptive mammographic image enhancement using first derivative and local statistics. *IEEE Transactions on Medical Imaging*, 16 :495–502, 1997.
- [KT97] D. G. Karakos and P. E. Trahanias. Generalized multichannel image-filtering structures. *IEEE Transactions on Image Processing*, 6 :1038–1045, 1997.
- [KUG09] E. Kalalembang, K. Usman, and I.P. Gunawan. Dct based local motion blur detection. *International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering*, 2009.
- [Lam96] C.J. Van Den Branden Lambrecht. Perceptual models and architectures for video coding applications. *PhD thesis, Ecole polytechnique Fédérale de Lausanne*, 1996.

- [LF80] G. Legge and J. Foley. Contrast masking in human vision. *Physics Education*, 70 :1458–4471, 1980.
- [Lil93] O. Lillesaeter. Complex contrast, a definition for structured targets and backgrounds. *Journal of the Optical Society of America*, 10 :2453–2457, 1993.
- [LLP98] Y.L. Lee, H.C. Lim, and H.W. Park. Blocking effect reduction of jpeg images by signal adaptative filtering. *IEEE Transactions Image Processing*, 7 :229–234, 1998.
- [Lub95] J. Lubin. A visual discrimination model for imaging system design and evaluation. *Vision model for target detection and recognition*, pages 245–283, 1995.
- [Lub97a] J. Lubin. A human vision system model for objective picture quality measurements. *International Broadcasting Convention*, 5 :498–503, 1997.
- [Lub97b] J. Lubin. Sarnoff jnd vision model : Algorithm description and testing. *Technical report, Sarnoff Corporation*, 1997.
- [Mar74] K.V. Mardia. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya Ser*, pages 115–128, 1974.
- [Mar80] S. Marcelja. Mathematical decription of the response of simple cortical cells. *Journal of the Optical society of America*, 70 :1297–1300, 1980.
- [Mar02] J-B. Martens. Multidimensional modeling of image quality. *Proceedings of IEEE*, 90 :133153, 2002.
- [Mat] Mathworks. Expectation maximization program. <http://www.mathworks.com/matlabcentral/fileexchange/10956>.
- [MDWE02] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi. A no-reference perceptual blur metric. *International Conference on Image Processing*, 2002.
- [MDWE04] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi. Perceptual blur and ringing metrics : application to jpeg2000. *Signal Processing : Image Communication*, 19 :163–172, 2004.

- [Mic27] A. Michelson. Studies in optic. *Chicago Press*, 1927.
- [MK05] R. Muijs and I. Kirenko. A no-reference blocking artifact measure for adaptive video processing. *European Signal Processing Conference*, 2005.
- [MMS04] R. Mantiuk, K. Myszkowski, and H.P. Seidel. Visible difference predictor for high dynamic range images. *International Conference on Systems, Man and Cybernetics*, 3 :2763–2769, 2004.
- [MS43] P. Moon and D. E. Spencer. The specification of foveal adaptation. *Journal of the Optical Society of America*, 33 :444–456, 1943.
- [MS44] P. Moon and D. E. Spencer. Visual data applied to lighting design. *Journal of the Optical Society of America*, 34 :230–240, 1944.
- [MS74] J. L. Mannos and D. J. Sakrison. The effects of a visual fidelity criterion on the encoding of images. *IEEE Transactions. Information Theory*, 4 :525–536, 1974.
- [MV93] T. Mitsa and K. Varkur. Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms. *International Conference on Acoustics, Speech, and Signal Processing*, pages 301–304, 1993.
- [MWYJ99] X. Marichal, M. Wei-Ying, and Z. Hong Jiang. Blur determination in the compressed domain using dct information. *International Conference on Image Processing*, 2 :386–390, 1999.
- [Nil85] N.B. Nill. A visual weighted cosine transform for image compression and quality assessment. *IEEE Transactions on communications*, 33 :551–556, 1985.
- [OHN98] S.H. Oguz, Y.H. Hu, , and T.Q. Nguyen. Image coding ringing artifact reduction using morphological post-filtering. *Workshop on Multimedia Signal Processing*, pages 628–633, 1998.
- [OKD09] T. Oelbaum, C. Keimel, and K. Diepold. Rule based no reference video quality evaluation using additionally coded videos. *IEEE Journal of Selected Topics in Signal Processing*, 3 :294 – 303, 2009.
- [OLL⁺03a] E. Ong, W. Lin, Z. Lu, X. Yang, S. Yao, F. Pan, L. Jiang, and F. Moschetti. A no reference quality metric for measuring image blur. *Inter-*

- national Symposium on Signal Processing and Its Applications*, 1 :469 – 472, 2003.
- [OLL⁺03b] E.P. Ong, W. Lin, Z.k. Lu, S. Yao, X.k. Yang, and L. Jinag. No-reference jpeg2000. *International Conference on Multimedia and Expo*, 1 :545–548, 2003.
- [OMM97] W. Osberger, A.J. Maeder, and D. McLean. A computational model of the human visual system for image quality assessment. *Digital Imaging Computing : Techniques Applications*, pages 337–342, 1997.
- [PAYG93] E. Peli, L.E. Arend, G.M. Young, and R.B. Goldstein. Contrast sensitivity to patch stimuli : effects of spatial bandwidth and temporal presentation. *Spatial Vision*, 7 :1–14, 1993.
- [PBHK97] N. Peter, J. Belhumeur, P. Hespanha, and D. Kriegman. Eigenfaces vs. fisherfaces : Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 711–720, 1997.
- [PCD06] Le Callet P., Viard-Gaudin C., and Barba D. A convolution neural network approach for objective video quality assessment. *IEEE Transactions on Neural Networks*, 17 :1316–1327, 2006.
- [PCL⁺08] N. Ponomarenko, M. Carli, V. Lukin, K. Egiazarian, J. Astola, and F. Battisti. Color image database for evaluation of image quality metrics. *International Workshop on Multimedia Signal Processing*, pages 403–408, 2008.
- [Pel90] Eli Peli. Contrast in complex images. *Journal of the Optical Society of America*, 7 :2032–2040, 1990.
- [PLR⁺04] F. Pan, X. Lin, S. Rahardja, W. Lin, E. Ong, S. Yao, Z. Lu, and X. Yang. A locally-adaptive algorithm for measuring blocking artifacts in images and videos. *International Symposium on Circuits and Systems*, 3 :925–928, 2004.
- [PLZ⁺09] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti. Tid2008 - a database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics*, 10 :30–45, 2009.

- [PSE⁺07] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin. On between-coefficient contrast masking of dct basis functions. *International Workshop on Video Processing and Quality Metrics*, 2007.
- [QDL10] A.Beghdadi Q.B. DO and M. Luong. A new adaptive image post-treatment for deblocking and deringing based on total variation method. *International Symposium on Signal Processing and Its Applications*, 2010.
- [Sak77] D.J. Sakrison. On the role of the observer and a distortion measure in image transmission. *IEEE Transactions on Communications*, 25 :1251–1267, 1977.
- [SAMB09] W. Soudene, K. Abed-Meraim, and A. Beghdadi. A new look to multichannel blind image deconvolution. *IEEE Transactions on Image Processing*, 18 :1487–1500, 2009.
- [SB06a] H.R. Sheikh and A.C Bovik. Image information and visual quality. *IEEE Transactions on Image Processing*, 15 :430–444, 2006.
- [SB06b] H.R. Sheikh and A.C Bovik. Image information and visual quality (in pixel domain). <http://live.ece.utexas.edu/research/quality/>, 2006.
- [SBC05] H.R. Sheikh, A.C. Bovik, , and L.K. Cormack. No-reference quality assessment using natural scene statistics : Jpeg2000. *IEEE Transactions on Image Processing*, 14 :1918–1927, 2005.
- [SBdV05] H.R. Sheikh, A.C. Bovik, and G. de Veciana. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, 14 :2117–2128, 2005.
- [SC10] A. Chetouani S. Chen, A. Beghdadi. A new color image quality index. *Workshop on Video Processing and Quality Metrics*, 2010.
- [SGE04] A. Shnayderman, A. Gusev, and A.M. Eskicioglu. A multidimensional image quality measure using singular value decomposition. *SPIE Image Quality and System Performance Conference*, 5294 :82–92, 2004.
- [SKH07] Z.M.P. Sazzad, Y. Kawayoke, and T. Horita. Spatial features based on no reference image quality assessment for jpeg2000. *International Conference on Image Processing*, 3 :517–520, 2007.

- [SkH08] Z.M. Parvez Sazzad, Y. kawayoke, and Y. Horita. Image quality evaluation database. http://mict.eng.u-toyama.ac.jp/database_toyama/, 2008.
- [SWCB06] H.R. Sheikh, Z. Wang, L. Cormack, and A.C. Bovik. Live image quality assessment database. <http://live.ece.utexas.edu/research/quality>, 2006.
- [SY09] X. Song and Y. Yang. A new no reference assessment metric of blocking artefacts based on hvs masking effect. *International Congress on Image and Signal Processing*, pages 1–6, 2009.
- [TCB07] S. Tourancheau, P. Le Callet, and D. Barba. Prediction of perceived quality differences between crt and lcd displays based on motion blur. *International Workshop on Video Processing and Quality Metrics*, 2007.
- [TH94] P.C. Teo and D. J. Hegger. Perceptual image distortion. *SPIE - Human Visual Processing and Digital Display*, 2179 :127–141, 1994.
- [TJZ96] Z. Tianxu, P. Jiaxiong, and L. Zongjie. An adaptive image segmentation method with visual nonlinearity characteristics. *IEEE Transactions on Systems, Man Cybern*, 26 :619–627, 1996.
- [TLHJZ04] H. Tong, M. Li, Z. Hong-Jiang, and C. Zhang. No reference quality assessment for jpeg2000 compressed images. *International Conference on Image Processing*, 5 :3539–3542, 2004.
- [TLZZ04] H. Tong, M. Li, Zhang, and C. Zhang. Blur detection for digital images using wavelet transform. *International Conference on Multimedia and Expo*, 1 :17–20, 2004.
- [TM01] D. Taubman and M. Marcellin. *JPEG2000 : Image compression fundamentals, standards and practice*. Kluwer Academic Publishers, 2001.
- [Tou09] S. Tourancheau. Caractérisation objective et modélisation psychovisuelle du flou de mouvement sur les écrans à cristaux liquides. impact sur la qualité perçue. *Rapport de thèse, Université de Nantes*, 2009.
- [VK08] S. Varadarajan and L.J. Karam. An improved perception-based no-reference objective image sharpness metric using iterative edge refine-

- ment. *International Conference on Image Processing*, pages 401–404, 2008.
- [VPM96] Bruce V., Green P., and Georgeson M. Visual perception : Physiology, psychology and ecology. *LEA*, page 110, 1996.
- [VQE00] VQEG. Final report from the video quality experts group on the validation of objective models of video quality assessment. *ftp ://vqeg.its.bldrdoc.gov/Documents/*, 2000.
- [Wat93] A. B. Watson. Dctune : A technique for visual optimization of dct quantization matrices for individual images. *Society for Information Display Digest of Technical Papers*, 14 :946–949, 1993.
- [WB02a] Z. Wang and A. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, 2 :81–84, 2002.
- [WB02b] Z. Wang and A. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, 2 :81–84, 2002.
- [WB09] Z. Wang and A. C. Bovik. Mean squared error : love it or leave it ? - a new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26 :98–117, 2009.
- [WBE00] Z. Wang, A.C. Bovik, and B.L. Evans. Blind measurement of blocking artefacts in images. *International Conference on Image Processing*, 3 :981–984, 2000.
- [WBE05] Z. Wang, A.C. Bovik, and E.P. Simoncelli E.P. Structural approaches to image quality assessment. *Handbook of image and video processing, 2nd Edition, Al Bovik, ed, Academic Press*, 2005.
- [WBSS04] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Smoncelli. Image quality assessment : from error visibility to structural similarity. *IEEE Transactions Image Processing*, 13 :600–612, 2004.
- [Wel68] W.T. Welford. Vector quantization. *IEEE Acoustics, Speech and Signal Processing Newsletter*, 3 :83–84, 1968.
- [WG84] H. R. Wilson and D. J. Gelb. Modied line-element theory for spatial-frequency and width discrimination. *Journal of the Optical society of America*, pages 124–131, 1984.

- [Wil98] M.H.F. Wilkinson. Optimizing edge detectors for robust automatic threshold selection : Coping with edge curvature and noise. *Graphical models and image processing*, 60 :385–401, 1998.
- [Win05] S. Winkler. *Digital Video Quality : Vision Models and Metrics*. John Wiley and Sons, 2005.
- [WLB95] S.J.P. Westen, R.L. Lagendijk, and J. Biemond. Perceptual image quality based on a multiple channel hvs model. *International Conference on Acoustics, Speech and Signal Processing*, pages 2351–2354, 1995.
- [WR05] H.R. Wu and K.R. Rao. *Digital Video Image Quality and Perceptual Coding*, volume ISBN : 0-8247-2777-0. CRC Press, 2005.
- [WS97] A.B. Watson and J.A. Solomon. Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America*, 14 :2379–2391, 1997.
- [WS05] Z. Wang and E. Simoncelli. Translation insensitive image similarity in complex wavelet domain. *International Conference on Acoustics, Speech and Signal Processing*, 2 :573–576, 2005.
- [WSB02] Z. Wang, H. Sheikh, and A. Bovik. No-reference perceptual quality assessment of jpeg compressed images, 2002.
- [WSB03] Z. Wang, E.P. Simoncelli, and A.C. Bovik. Multi-scale structural similarity for image quality assessment. *Asilomar Conference on Signals, Systems and Computers*, 2003.
- [WSH04] Z. Wang, E.P. Simoncelli, and H. Hughes. Local phase coherence and the perception of blur. *Advances in Neural Information Processing Systems*, 16 :786–792, 2004.
- [WTLS08] X. Wang, B. Tian, C. Liang, and D. Shi. Blind image quality assessment for measuring image blur. *Congress on Image and Signal Processing*, 1 :467–470, 2008.
- [WX09] S. Wu and S. Xie. An efficient blind method for image quality measurement. *International Conference on Information Communications and Signal Processing*, pages 1–5, 2009.

- [WYSV97] A.B. Watson, G.Y. Yang, J.A. Solomon, and J Villaseñor. Visibility of wavelet quantisation noise. *IEEE Transactions on Image Processing*, 6 :1164–1175, 1997.
- [YW98] M. Yuen and H. R. Wu. A survey of hybrid mc/dpcm/dct video coding distortions. *Signal Processing*, 70 :247–278, 1998.
- [YWL05] Y. You, A. Wang, and F. Lu. A wavelet approach to image quality assessment using neural networks. *EE 371D, Electrical and Computer Engineering, University of Texas at Austin*, 17 :1316–1327, 2005.
- [Zen99] B. Zeng. Reduction of blocking effect in dct-coded images using zero-masking techniques. *Signal Processing*, 79 :205–211, 1999.
- [ZL10] J. Zhang and T.M. Le. A new no reference quality metric for jpeg2000 images. *IEEE Transactions on Consumer Electronics*, 56 :743–750, 2010.
- [ZXL08] J. Zhou, B. Xiao, and Q. Li. A no reference image quality assessment method for jpeg2000. *World Congress on Computational Intelligence*, pages 863–868, 2008.
- [ZZYL08] G. Zhai, W. Zhang, X. Yang, and W. Lin. Efficient image deblocking based on post filtering in shifted windows. *IEEE Transactions on Circuits and Systems for Video Technology*, 18 :122–126, 2008.